ESTIMATION OF IMPEDANCE USING SEISMIC REFLECTION DATA BASED ON TRANSDIMENSIONAL INVERSION

A Dissertation

by

SIREESH KUMAR DADI

Submitted to the Office of Graduate and Professional Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Chair of Committee,	Richard Gibson
Committee Members,	Benchun Duan
	Michael King
	Yuefeng Sun
Head of Department,	Rick Giardino

May 2014

Major Subject: Geophysics

ABSTRACT

Deterministic seismic inversion can obtain optimal results when there is a linear relationship between data and model parameters during local optimization (single optimum solution). For nonlinear geophysical problems and in the presence of multiple local minima for a cost function, global optimization techniques are necessary to characterize the global minimum solution. Stochastic, model-based seismic inversion is a widely used global optimization technique and Markov Chain Monte Carlo (MCMC) method is a natural choice to sample model parameters during the random walk. In this dissertation, I apply a sampling technique called reversible jump Markov Chain Monte Carlo (rjMCMC) to traverse the model space. A key property of this approach is that it automatically changes the layer thicknesses and number of layers, thereby predicting the optimum number of model parameters during inversion. The method applies Bayesian inversion, with rjMCMC sampling, so that it also quantifies the uncertainty in model parameters based on an ensemble of models. I apply Bayesian inversion with rjMCMC sampling for two applications. In the first application, I define upscaling velocity logs as an inversion problem to obtain optimal models and quantify uncertainty of upscaled models at the well location. The upscaled velocity at the well locations can be subsequently used to stabilize velocity inversion during Full waveform Inversion (FWI) for seismic imaging purposes. In the second application, I perform post-stack seismic inversion to obtain shallow impedance structure of the TAMU and Ori volcanoes at the Shatsky Rise oceanic plateau. Since impedance is a rock property, it is used to discriminate basalt rock types, which gives insight into the late-stage evolution of both the volcanoes.

To my beloved grand parents

ACKNOWLEDGEMENTS

As I walk across the Texas A&M campus everyday, I feel extremely thankful to God for giving me this opportunity to pursue doctoral studies here. I would like to thank Dr. Gibson for offering practical solutions whenever I hit a roadblock in my research. I am thankful for his support to define scientific problems and execute the tasks with great focus. Dr Gibson has been an inspiring teacher; I learned a tremendous amount from him in classroom and during regular research interactions. Along this journey, I feel fortunate to have worked with Dr. Yuefeng Sun during the Imperial Barrel Award (IBA) Competition, which later played a vital role in shaping my career choice. The days we spent working on the IBA project were among the most memorable days of my stay at Texas A&M.

I would like to thank Dr. Michael King and Dr. Benchun Duan for their guidance in classroom and suggestions on improving my research approach. Dr. King gave a thorough overview of upscaling geologic models, which is an integral part of this dissertation. The coursework by Dr. Pope, Dr. Ahr and Dr Wiltschko is very insightful and assisted in a smooth transition from an engineering background to geosciences. I am grateful to Dr. David Dahl for teaching the significance of Bayesian methods, which laid a solid foundation for my Ph.D research. I am also thankful to my research group members Kai Gao, Au, Muhammed, Jungrok, Dehan, and Illaan for their insights on my research work during the team meetings. I would like to extend a special thanks to Kai Gao who always assisted me with general questions on various softwares I used during my doctoral studies.

My research is highly interdisciplinary, and I am grateful to my collaborators for appreciating my research problems, specifically, my collaboration with Kainan Wang

helped me greatly improve the dissertation and implement the rigor of statistical approaches. I would like to thank Dr. Zhang and Dr. Will Sager for providing unconditional access to the seismic reflection data from Shatsky Rise, which helped me focus on the inversion method.

My memories at Texas A&M will forever be remembered by the people I met here. In the last year, Sandeep, Sarath, Sujan, and Krishna have been part of our social group whose kindness and generosity have been unmatched. On a longer timescale during my Ph.D, the days I spent with Abhishek, Ahmed, Bharath, Chaitanya, Hamid, Hayes, John, Karthik, Maruti, Sandeep, Sharbel, Tejasvi, and Vivek were filled with lots of adventures and fun conversations.

My Ph.D journey had many ups and downs and would not have been possible without the unconditional love and support from my family. I would like to thank my parents for their confidence in my ability to take career choices of my interest. My sister, Swathi, was always there for me in trying times. I am grateful to her for taking care of my parents when they needed me. My uncle, Tanakala Prasad, has been a great inspiration for me and we have lived this dream together. He identified the potential in me and encouraged me to aim high in my academics. Finally, I dedicate this dissertation to my late grand parents, Seethayyamma and Rambabu for their immense love and blessings.

TABLE OF CONTENTS

Page

LIST OF FIGURES

FIGURE Page

LIST OF TABLES

TABLE Page

1. INTRODUCTION

1.1 Overview

An inverse problem can be set up in a Bayesian framework by obtaining the posterior probability distribution (ppd) of the model parameters of interest but demands prior information and a likelihood function, sometimes referred to as as a fitness or error function (Hoff, 2009). Bayesian inversion has been applied to many geophysical problems, including lithology and fluid prediction (Ulrych et al., 2001; Rimstad and Omre, 2010; Rimstad et al., 2010; Ulvmoen and Omre, 2010), the integration of seismic inversion with rock physics (Grana and Della Rossa, 2010), Amplitude variation with Offset (AVO) inversion (Buland and Omre, 2003), time lapse inversion (Buland and El Ouair, 2006) and inversion to determine geological models (Cardiff and Kitanidis, 2009). All the above mentioned work uses analytical solutions for the Bayesian formulation, applying assumptions such as the existence of a conjugate prior and that the ppd is approximated to Gaussian or another form of analytical kernel distributions. In cases when an analytical solution for ppd cannot be calculated, the Bayesian framework can solve for ppd using Markov Chain Monte Carlo (MCMC) (Mosegaard and Tarantola, 1995; Eidsvik et al., 2004; Hong and Sen, 2009; Malinverno, 2002; Gibson and Hwang, 2009; Gunning and Glinsky, 2004; Sambridge, 1999) techniques. MCMC techniques will assist in better sampling the model space, however, the procedure remains to be computationally intensive and requires more efficiency.

In general, Monte Carlo-based geophysical inversions face two problems: the first is to generate a sampling technique which can effectively sample the model space and improve mixing properties and the second is to build rapid forward models.

Multiscale MCMC (MSMCMC) is a multilevel (multi chain) sampling technique designed to address the former of the two problems that can transfer information between different scales during the process of accepting/rejecting the samples, thereby expanding the sampled space for all the scales. Hong and Sen (2009) demonstrated that using Multiscale MCMC has reduced the computational cost of seismic inversion by 40% while obtaining finer scale reservoir properties (i.e layer thickness of 8-10 m) compared to fixed scale MCMC.

In this dissertation, I use the reversible jump MCMC sampling technique, which allows for change in dimensions of model parameters, thereby allowing intrachain transfer of information between different dimensional models. Reversible jump MCMC considers the number of unknowns itself as an unknown, and as a result the number of model parameters changes during the random sampling of potential models. Due to this property of allowing the number of model parameters to change, rjMCMC is often referred to as a transdimensional sampling technique. Reversible jump MCMC was introduced by Green (1995) and Green (2003) and first adopted in the geophysical literature by Malinverno (2002) to invert for DC resistivity. When implementing Bayesian inversion with rjMCMC sampling, Malinverno (2002) showed that simple models will be preferred to complex models during the random sweep. Malinverno and Leaney (2005) inverted a vertical seismic profile (VSP), using rjMCMC, to obtain p-wave velocity, s-wave velocity and density. They found the uncertainty of inverted p-wave velocity is lower when using VSP as reference data compared to the case when they used zero-offset data. Malinverno and Leaney (2005) used a linearized forward model which generates only primary reflections.

Reversible jump MCMC is used to obtain shear wave velocity in crustal scale studies by inverting receiver functions and surface dispersion data (Agostinetti and Malinverno, 2010; Sambridge et al., 2013). Dettmer et al. (2013) inverted seabed

reflection data to obtain velocity dispersion and frequency dependent attenuation of shallow sea sediments and their uncertainties. Transdimensional inversion is also used to identify abrupt changes in geochemical records (e.g., total carbon, $.\delta^{13}\mathrm{C}$) as a function of depth or time due to climate-variation, to calculate the number of such changes, and to measure mean values between change points (Gallagher et al., 2011). A similar analysis is done in this dissertation, and by Reading and Gallagher (2013), where abrupt and gradual changes in well logs are identified as probability distributions. The objective of the change-point analysis done by Reading and Gallagher (2013) is to identify lithology contrast depths and to calculate variation of well log values with in each lithology. Reading and Gallagher (2013) averaged well logs within each layer using arithmetic averaging, and calculated the cost function as a mismatch between averaged log and log data. While Reading and Gallagher (2013) used arithmetic averaging, Gibson and Hwang (2009) used Backus averaging for velocity log upscaling and the cost function is the mismatch between seismic signal from fine layered media and upscaled medium. Reversible jump MCMC algorithms are also used for 2-D inversion problems with various model parameterization schemes (spatial discretization) for seismic tomography and flow in porous media applications (Mondal et al., 2010; Bodin and Sambridge, 2009; JafarGandomi and Binley, 2013).

I implemented rjMCMC sampling in Bayesian inversion framework for two new applications in this dissertation. In the first application, I define upscaling velocity log as a transdimensional inversion problem for the first time to obtain optimal upscaled models, and to estimate the uncertainty of upscaled velocity logs. The results suggest optimal models obtained using transdimensional inversion can reproduce seismic propagation from fine scale media better than Backus averaging methods. In the second application, I invert seismic reflection data on Shatsky Rise oceanic plateau to obtain shallow impedance structure, which is used to study late-stage evolution

of two major volcanoes at Shatsky Rise. The shallow impedance at TAMU Massif, which is the largest volcano at Shatsky Rise, show there is more localized explosive volcanism at the summit compared to that at the flank. Below, I summarize the motivation and objectives for each of the applications.

1.2 Upscaling velocity logs

1.2.1 Motivation

The motivation to generate coarse-scale layered models is for practical reasons including realistic computation cost to build forward models. Full wave forward modeling of fine scale heterogeneities is computationally challenging especially for 2-D and 3-D models and demands the need for an accurate effective medium representation. In a grid-based wave modeling method, the computational cost is proportional to the number of grids given by

$$N \propto N_x N_y N_z N_t, \tag{1.1}$$

where N_x, N_y, N_z and N_t are the number of grids in the x-, y-, z-directions and time steps, respectively. The computational time for full wave modeling, when the fine grid is upscaled to a coarser grid with grid ratio of 2, decreases the computation by 8 times for 2-D and 16 times for 3-D.

Upscaling heterogeneities in 2-D and 3-D geologic medium without changing the seismic wave field still remains to be a challenging problem for geophysicists, however, in the case of 1-D upscaling, effective medium approximations like Backus averaging are widely accepted. A stochastic approach for upscaling well logs can quantify the uncertainty in velocity log upscaling and also quantify the difference between the actual effective medium elastic properties and the Backus average values. Grechka (2003) quantified the error between window based Backus average velocity and velocity obtained by numerical modeling. Grechka (2003) suggested the lower values

of averaged elastic moduli obtained from Backus average can be explained by the nature of the isostress averaging scheme which is a density weighted harmonic mean of velocity.

1.2.2 Objectives

In chapter II, I compare well log upscaling using a simple MCMC algorithm, with a constant number of layers, to rjMCMC results where the number of layers is treated as an unknown. The two approaches were applied to two different depth intervals from a well log from the North Sea. I improved the Backus averaging methodology used by Gibson and Hwang (2009), which allows only boundaries to move during the sampling, by applying three potential moves, or changes, from one iteration to the next: changing the boundary depths, adding new boundaries or deleting boundaries. In all cases, I also added a random perturbation of the velocity from Backus average values in all the three moves. The inversion is performed in a Bayesian framework using rjMCMC sampling technique and I used simulated annealing (rjMCMCSA) to guide the global optimization. The advantage with the current averaging technique is that both sharp contacts like unconformities and gradation contacts are preserved, unlike the Sequential Backus averaging where only gradational contacts are preserved. None of the existing effective medium theories except numerical modeling by Grechka (2003) and the present study calculate the deviation of optimal velocities from layer-based Backus averaged values. and to derive petrophysical properties from the inverted coarse-scale elastic properties which can be used for fluid flow modeling.

In chapter III, I modified the rjMCMCSA upscaling approach, which generates only optimal upscaled velocity models, by sampling models in the full posterior distribution that generate a true estimate of uncertainty. The results obtained when

using SA are obtained by tempering the likelihood function which allows the sampler to rapidly converge to the modes of the full posterior distribution. The results obtained with straight rjMCMC without tempering the likelihood function generate a true estimate of the uncertainty of the upscaled velocity models, but estimating the optimal model requires sampling many more models, and hence more forward model evaluations are needed. I compare the average and estimates of uncertainty from upscaled models for two well log intervals from a well in North Sea with and without tempering the likelihood function. I extended the upscaling method by using multi-offset seismograms as reference seismograms to investigate the effect of using multi-offset seismograms on upscaled models. I chose well logs measured at location U1347 on the Shatsky Rise oceanic plateau, which has basalt rocks interbedded with sediments (high impedance contrasts) to compare the estimates of uncertainty of upscaled models when using zero-offset and multi-offset seismograms as reference seismograms. I introduced an empirical convergence assessment for RJMCMC sampler to determine a point in the random sweep where the sampler had ‘‘converged.” Convergence here refers to the point after which the samples are generated from the full posterior distribution.

1.3 Seismic inversion of Shatsky Rise

1.3.1 Motivation

Shatsky Rise is an oceanic plateau located in the northwest Pacific Ocean which is formed as a result of massive eruption of lavas. Shatsky Rise is identified as one of its kind among all the oceanic plateaus on Earth because of its large size and it remained subaqueous since its inception. Shatsky Rise is formed at a triple junction during late Jurassic and the duration for the formation of the entire plateau is believed to be around ∼21 million years (Zhang et al., 2012). The Shatsky Rise oceanic plateau

consists of three volcanoes, the TAMU Massif, Ori Massif and the Shirsov Massif. The TAMU Massif is the oldest of the three volcanoes and Shirsov Massif is the earliest. The oceanic plateau formed as a result of initial high rates of lava effusion during the late Jurassic resulting in the formation of TAMU Massif which is one of largest volcanoes on Earth. The Ori and Shirsov massifs are formed at a later stage during a phase of waning effusion rate of lava flows. Massive sheet like basalts are formed due to high lava effusion rate which are identified as continuous seismic reflections from the summit to the flanks of TAMU massif (Zhang et al., 2012). Pillow basalts are formed due to localized explosive volcanism and are often intermingled with volcanoclastic sediments as observed in drilling cores.

The thickness of basalt and sediment layers at Shatsky Rise vary from a few meters to tens of meters and this variation in thickness poses a challenge to perform fixed dimensional stochastic inversion to obtain impedance values. For example, when the thinnest layer thickness is 20 m in a 1 km interval, we need to define at least 50 model parameters during inversion. To avoid this problem, I use a transdimensional approach for seismic inversion where the number of layers, layer boundaries and impedance values are defined as model parameters. Using the reversible jump MCMC approach reduces the computational cost and avoids curse of dimensionality which is noticed in high-dimensional inversion problems.

1.3.2 Objectives

The objective of this study is to perform impedance inversion from 2D post-stack seismic reflection data collected during IODP Expedition 324 in 2012. I study the shallow impedance structure at Shatsky Rise and draw inference about the late-stage evolution of the TAMU and Ori Volcanoes in the Shatsky Rise oceanic plateau. Zhang et al. (2012) interpreted the intra basalt reflections in the seismic

reflection data as impedance contrast between different basalt types and between basalt and volcanoclastic sediments. I inverted seismic reflection data to obtain impedance as it is a rock property and can be used to characterize different basalt rock types. I perform waveform based Bayesian inversion instead of using only travel time picks to accurately capture the uncertainty of model parameters when the thickness of layers is less than tuning thickness. I modified the trandimensional inversion approach by Malinverno and Leaney (2005), which used a linearized forward model that generates only primary reflections, by using a non-linear forward model (Propagator matrix method). The linear forward model will be a reasonable assumption only when the contrasts in elastic properties are small, however, the contrast between basalts and sediments is too large to ignore internal multiples.

1.4 Dissertation structure

In this section, I give a brief overview of all the chapter and the structure of the dissertation. Chapter I explains the motivation for choosing transdimensional inversion and performing well log upscaling. Chapter II explains the advantages of using rjMCMCSA while performing velocity log upscaling and determines the deviation of the optimal velocities from Backus average values while using normal incidence seismogram as reference seismogram. Chapter III compares the upscaled velocity results while using rjMCMC sampling without SA, with upscaled velocities obtained while using rjMCMCSA. Chapter III also compares upscaled velocity results when using normal incidence and multi-offset seismograms as reference seismograms. In Chapter IV, I use transdimensional inference to invert for impedance structure of basalts at TAMU and Ori volcanoes which are part of the Shatsky Rise oceanic plateau. The data used for seismic inversion include 2-D Multi-channel seismic (MCS) reflection lines collected on the Shatsky Rise oceanic plateau by the R/V

Marcus G. Langseth cruise in 2012. Chapter V concludes the main findings in the dissertation and Chapter VI is an appendix which includes a detailed description of the acceptance criteria while using rjMCMC sampling.

2. VELOCITY LOG UPSCALING BASED ON REVERSIBLE JUMP MCMC SIMULATED ANNEALING

2.1 Introduction

Well log upscaling is a traditional practice to compare elastic properties measured at higher frequencies to those obtained from lower frequencies measurements, such as surface seismic or VSP data. Simple statistical methods (arithmetic, harmonic or geometric averages) or analytical averaging techniques like Backus averaging are used for well log upscaling. Backus averaging approximates the equivalent medium as a transversely isotropic (TI) medium which can be defined with five independent elastic moduli C_{11}, C_{33}, C_{44}, C_{66} and C_{13}. The vertical velocities in a specific layer can be obtained from Backus averaged elastic moduli C_{33}, C_{44} and density. Backus averaging can be performed as layer based upscaling (Folstad and Schoenberg, 1992; Prüssmann, 1996; Gibson and Hwang, 2009) and smooth window based upscaling (Rio et al., 1996; Liner and Fei, 2006; Lindsay and Van Koughnet, 2001; Sayers, 1998; Tiwary et al., 2007).

Backus average has been implemented successfully for many real well log examples; however the technique has its own limitations depending on how the averaging is performed and due to the inherent assumptions in the Backus average derivation. Using smooth window based upscaling can preserve gradational contacts (Lindsay and Van Koughnet, 2001) but sharp contacts like unconformities are smoothed over. The ambiguities in layer-based Backus averaging for a well log include the number of layers to be considered and whether the thickness of the layers should be uniform or non uniform. Applications of window-based Backus averaging often find that the choice of a specific window length for upscaling, to equivalent Transverse Isotropic

(TI) medium, remains uncertain. For example, Liner and Fei (2006), Folstad and Schoenberg (1992) and Avseth et al. (2005) suggested window lengths of $\lambda/4$, $\lambda/3$ and $\lambda/10$ respectively for their studies. The different window lengths picked by each of the authors depends on the acceptable tolerance of the match between the propagating wave signal through the equivalent medium compared to that through the fine layered medium. Liner and Fei (2006) used window-based Backus average for upscaling and identified non zero values of anisotropy values at sharp contacts which might be an artifact of using smooth window based Backus averaging. Lindsay and Van Koughnet (2001) introduced Sequential Backus averaging which uses a dynamic upscaling window dependent on the wavelength. The upscaling window changes with depth depending on the velocity and frequency.

In some cases, lateral heterogeneity will cause inaccuracies in upscaling, and Sayers (1998) derived the analytical expressions for average elastic properties of an equivalent medium when the medium includes elliptical inclusions (characterized by volume fraction of the inclusion and with an aspect ratio). This representation of a medium is more realistic in turbidite reservoirs where shale lens which act as baffles or barriers exist in channel sands. Grechka (2003) deliberately studied two cases, first when including fractures in the laterally continuous fine layers and second by including random inclusions, to show elastic moduli obtained from traditional Backus averaging deviate from average values obtained from numerical modeling. The traditional Backus averaging takes into account the interaction between fine layers but not the interactions between the fine layers and fractures. Grechka (2003) followed a numerical approach for calculating complex stresses and strains and averaged them in a volume to obtain the effective medium parameters. The averaged stresses and strains are substituted in the Hooke's law and elastic moduli are calculated. The Backus averaged values consistently underestimates equivalent medium elastic

moduli obtained from numerical modeling under static loading (frequency = 0). The lower values of averaged elastic moduli obtained from Backus average can be explained by the nature of the averaging scheme which is either harmonic or weighted harmonic. Tiwary et al. (2007) introduced an averaging scheme with pair and multi-correlation functions which account for scattering in the medium through elastodynamic Green's functions between correlation points which in turn depend on the medium properties. Gibson and Hwang (2009) followed a stochastic approach of layer-based Backus averaging on a test well with a fixed number of layers but allowed all the boundaries of the layers to change. They used a Markov Chain Monte Carlo approach to determine a model with a minimal difference between the seismic response of the upscaled model and the original well log, computing the exact 1-D waveform using the propagator matrix method (Aki and Richards, 2002). An advantage of this stochastic approach is that it produces probability distributions for depths of layer boundaries, giving quantitative insights into uncertainty in the upscaled model.

In this chapter, I demonstrate the effectiveness of rjMCMC by comparing well log upscaling using a simple MCMC algorithm with a constant number of layers to rjMCMC results where the number of layers is treated as an unknown. The two approaches were applied to two different depth intervals from a well log from the North Sea. I thus improve the Backus averaging methodology used by Gibson and Hwang (2009), which allows only boundaries to move during the sampling, by applying three potential moves, or changes, from one iteration to the next: changing the boundary depths, adding new boundaries or deleting boundaries. In all cases, I also added a random perturbation of the velocity from Backus average values in all the three moves. The inversion is performed in a Bayesian framework using rjMCMC sampling technique and I used simulated annealing to guide the global

optimization. Samples are accepted and rejected based on a seismic likelihood function which determines the match between synthetic seismogram from upscaled model and reference seismogram obtained from well log. Our methodology addresses the two ambiguities while performing Backus averaging by allowing the number of layers itself as unknowns and the boundaries to change. The advantage with the current averaging technique is, both the sharp contacts like unconformities and gradation contacts are preserved unlike the Sequential Backus averaging where only gradational contacts are preserved. None of the existing effective medium theories except numerical modeling by Grechka (2003) and the present study calculate the deviation of optimal velocities from layer-based Backus averaged values.

Below I first summarize the key results from Bayesian inverse theory and then describe the rjMCMC method and the key concepts that make its application different from conventional MCMC implementations. I also summarize the application to the well log upscaling approach, and present results for its application to the depth intervals from the North Sea well log. These tests show that the method can find a model with an optimal number of layers, and that the corresponding misfit of the predicted synthetic seismograms from those computed for the original well logs is smaller than for alternate, conventional upscaling approaches.

2.2 Theory

2.2.1 Bayes Theorem

An inverse problem can be set up in a Bayesian framework by obtaining the posterior probability distribution (ppd) of the model parameters of interest but demands prior information and a likelihood function, sometimes referred to as a fitness or error function (Hoff, 2009). Bayesian inversion has been applied to many geophysical problems, including lithology and fluid prediction (Ulrych et al., 2001;

Rimstad and Omre, 2010; Rimstad et al., 2010; Ulvmoen and Omre, 2010), the integration of seismic inversion with rock physics (Grana and Della Rossa, 2010), AVO inversion (Buland and Omre, 2003), time lapse inversion (Buland and El Ouair, 2006) and inversion to determine geological models (Cardiff and Kitanidis, 2009). All the above mentioned work uses analytical solutions for Bayesian formulation with assumptions like a conjugate prior exist and the ppd is approximated to Gaussian or other form of analytical kernel distributions.

Bayes theorem updates our belief on the prior information based on observed data (Hoff, 2009). The posterior probability distribution of the model parameters given the observed data can be expressed as

$$P(\mathbf{m}|\mathbf{d}) = \frac{P(\mathbf{d}|\mathbf{m})P(\mathbf{m})}{P(\mathbf{d})}. \tag{2.1}$$

In equation (2.1) $\mathbf{m}$ is the model parameter vector given by $(k, \mathbf{z}, \mathbf{v})$ where k is the model dimension (i.e number of layers), $\mathbf{z}$ is the vector of boundary depths and $\mathbf{v}$ is the vector of layer velocities. The model dimension parameter k is considered an unknown while solving for $(\mathbf{z}, \mathbf{v})$ resulting in a transdimensional approach. $\mathbf{d}$ is the reference data used to calculate the seismic likelihood, and the denominator $P(d)$ is the probability distribution of the reference data, for all possible models, which is a complex summation given by $\sum_m P(d|m)P(m)$. Due to complex parametrization, multidimensionality, multimodal nature of the prior reservoir properties (prior distribution) and nonlinear nature of the forward problem (likelihood function), a direct analytical approximation for the ppd is not possible in general. However, adopting a Markov Chain Monte Carlo sampling technique using a Metropolis-Hastings algorithm eliminates the need to calculate the total evidence while facilitating the quantification of the earth models during inversion.

2.2.2 Likelihood distribution

The likelihood function $P(\mathbf{d}|\mathbf{m})$ expresses the probability of the occurrence of the reference data given the elastic model. In inversion applications, the likelihood function may be computed in several ways, such as the RMS error of the predicted and reference seismograms (Gibson and Hwang, 2009), a cross correlation (Sen and Stoffa, 1991), and Shannon's entropy (JafarGandomi and Binley, 2013). In this work, I used the RMS error approach to calculate likelihood function, assuming a multivariate Gaussian distribution for data error. The likelihood function is then given by

$$P(\mathbf{d}|\mathbf{m}) = \frac{1}{(2\pi)^{(n/2)}|\mathbf{C_d}|^{(1/2)}} \exp[-\frac{1}{2}(g(\mathbf{m}) - \mathbf{d})^T \mathbf{C_d}^{-1}(g(\mathbf{m}) - \mathbf{d})], \tag{2.2}$$

where $\mathbf{C_d}$ is the data covariance error matrix, and $\mathbf{g}(\mathbf{m})$ is the forward model function representing the synthetic seismogram for upscaled model $\mathbf{m}$. One of the challenges in geophysical inversion has been when using multiple datasets. The data types have varying sensitivities to model parameters and different levels of noise.

2.2.3 Reversible jump MCMC algorithm

2.2.3.1 Acceptance criteria for MCMC algorithm

The idea of using a sampling technique like MCMC with a Bayesian formulation is to generate samples from the ppd which describe the most likely model parameters which can fit the data and also estimate uncertainty based on the ensemble of models. A Markov Chain approach is an iterative sampling technique where a model in iteration i is dependent only on the previous iteration $i-1$. The criteria for model

acceptance is based on the Metropolis-Hastings (M-H) ratio (Hastings, 1970)

$$\alpha(\mathbf{m}', \mathbf{m}) = min[1, \textbf{prior ratio} * \textbf{likelihood ratio} * \textbf{proposal ratio}], \tag{2.3}$$

The Metropolis-Hastings ratio determines whether a new model is accepted or rejected. A value of α greater than one implies better match between model and reference seismogram and thereby the new model is always accepted; however, when α is less than one, instead of rejecting the model, it is accepted with a probability. Accepting the new models with a probability assists in sampling the whole model space and avoids stagnation at a local optimum. The prior probability is the estimated distribution of model parameters, in our case often defined based on rock physics depth trends. A proper prior will avoid sampling unrealistic model parameter values. The likelihood ratio compares the seismic response between two transition models, and the proposal distribution selects the next move in a random walk. The proposal distribution in this study is a Gaussian distribution with mean value equal to the model parameter at iteration i. The difference in model parameters from one iteration to the next that is allowed by the proposal distribution is an important parameter to adjust properly, as it controls the acceptance ratio and the number of iterations required to achieve a stationary distribution of model parameters.

2.2.3.2 Acceptance criteria for rjMCMC algorithm

When using direct MCMC sampling, the model parameter $\mathbf{m}$ is a vector given by $(\mathbf{z}, \mathbf{v})$, whereas rjMCMC allows the number of parameters also to change during random walk. In our case, for example, I allow the number of layers to vary from one iteration to the next.

The M-H ratio for the reversible jump MCMC formulation can be expressed as

$$\alpha(\mathbf{m}', \mathbf{m}) = min[1, \frac{p(\mathbf{m}')}{p(\mathbf{m})} \frac{p(\mathbf{d}_{obs}|\mathbf{m}')}{p(\mathbf{d}_{obs}|\mathbf{m})} \frac{q(\mathbf{m}|\mathbf{m}')}{q(\mathbf{m}'|\mathbf{m})} |\mathbf{J}|], \tag{2.4}$$

where the model at $i+1$, $\mathbf{m}_{i+1}= \mathbf{m}'$ is accepted with probability $\alpha(\mathbf{m}', \mathbf{m})$, and $\mathbf{m}_{i+1}= \mathbf{m}$ is accepted with a probability 1-$\alpha(\mathbf{m}', \mathbf{m})$. $q(\mathbf{m}'|\mathbf{m})$ is the proposal distribution which determines the jump from $\mathbf{m}$ to $\mathbf{m}'$. In this study, $q(\mathbf{m}'|\mathbf{m})$ is a Gaussian distribution with mean value at $\mathbf{m}$.

2.2.3.3 Jacobian

$|\mathbf{J}|$ is the determinant of the Jacobian which normalizes the difference between volume of different dimensions (Green, 2003). $|\mathbf{J}|$ is calculated only when the sampling algorithm allows change in dimension of the model space during the random sweep. However, it turns out $|\mathbf{J}|$ is equal to 1 for the style of rjMCMC implemented in this dissertation while the dimension of the model space changes. Below I discuss how $|\mathbf{J}|$ is equal to 1 when increasing or decreasing the dimensions of the model space. In the case when the dimensions increase during an iteration in the Markov Chain, the transformation of the model parameters can be represented as

$$\mathbf{m} = (k, \mathbf{z}, \Delta z_i, \mathbf{v}, \Delta v_i) \longleftrightarrow (k', \mathbf{z}, z_{i+1}, \mathbf{v}, v_{i+1}) = \mathbf{m}', \tag{2.5}$$

where Δz_i and Δv_i are the random perturbations allowed for layer boundary at z_i and velocity of layer i respectively. k' and k are the dimensions of the model space in the proposed and current iterations respectively. z_{i+1} and v_{i+1} are the layer boundary depth and velocity in layer i in the proposed iteration after perturbation and are calculated as follows

$$z_{i+1} = z_i + \Delta z_i, \tag{2.6}$$

$$v_{i+1} = v_i + \Delta v_i. \tag{2.7}$$

Since in a given iteration only i^{th} layer boundary or velocity are allowed to perturb, the vectors $\mathbf{z}$ and $\mathbf{v}$ in equation 2.5 can be replaced with z_i and v_i. After making the substitution for the layer boundary and velocity vectors the transformation shown in equation 2.5 can be written as

$$\mathbf{m} = (k, z_i, \Delta z_i, v_i, \Delta v_i) \longleftrightarrow (k', z_i, z_{i+1}, v_i, v_{i+1}) = \mathbf{m}'. \tag{2.8}$$

Note in the above formulation only k is a discrete model parameter, but both $\mathbf{z}$ and $\mathbf{v}$ traverse in a continuous model space. Denison et al. (2002) showed the Jacobian determinant for discrete transformations is always 1. Hence the Jacobian matrix determinant for the transformation of continuous variables when the dimension space is increasing is written as

$$|\mathbf{J}| = \left| \frac{\delta(z_i, z_{i+1}, v_i, v_{i+1})}{\delta(z_i, \Delta z_i, v_i, \Delta v_i)} \right|. \tag{2.9}$$

Using equations 2.6 and 2.7, the determinant of $\mathbf{J}$ can be expanded as below

$$|\mathbf{J}| = \begin{vmatrix} \frac{\delta z_i}{\delta z_i} & \frac{\delta z_i}{\delta(\Delta z_i)} & \frac{\delta z_i}{\delta v_i} & \frac{\delta z_i}{\delta(\Delta v_i)} \\ \frac{\delta z_{i+1}}{\delta z_i} & \frac{\delta z_{i+1}}{\delta(\Delta z_i)} & \frac{\delta z_{i+1}}{\delta v_i} & \frac{\delta z_{i+1}}{\delta(\Delta v_i)} \\ \frac{\delta v_i}{\delta z_i} & \frac{\delta v_i}{\delta z_i} & \frac{\delta v_i}{\delta z_i} & \frac{\delta v_i}{\delta z_i} \\ \frac{\delta v_{i+1}}{\delta z_i} & \frac{\delta v_{i+1}}{\delta z_i} & \frac{\delta v_{i+1}}{\delta z_i} & \frac{\delta v_{i+1}}{\delta z_i} \end{vmatrix}$$

Simplifying the above equation for $|\mathbf{J}|$ will result in

$$|\mathbf{J}| = \begin{vmatrix} 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 1 \end{vmatrix} = 1$$

Bodin et al. (2012) mentioned the value of $|\mathbf{J}|$ when the dimensional space is decreasing is given by $|\mathbf{J}^{-1}|_{birth}$. Hence $|\mathbf{J}|_{death}$=1. The determinant of Jacobian conveniently turned out to be 1 in both the cases when the dimensional space is increasing and decreasing thereby simplifying the acceptance ratio for rjMCMC.

Notice the probabilities in the numerator of equation (2.4) are defined for $\mathbf{m}'$ except the proposal distribution for $\mathbf{m}'$ conditional to $\mathbf{m}$ ($q(\mathbf{m}'|\mathbf{m})$) is in the denominator; this is because, in the absence of proposal ratio in the M-H ratio, when a random walk reaches a local minimum, the subsequent jumps will be highly discouraged by the likelihood ratio as the fit becomes worse. To avoid the trapping of random walk in a local optimum, new models away from the local minimum need to be accepted. Introducing a proposal distribution in the denominator improves the acceptance rate as the proposal probability is low for proposed models away from the local optimum.

2.2.3.4 Detailed balance condition

It is essential to show that the transition kernel generated by the Markov Chain during rjMCMC sampling satisfies detailed balanced condition iteration-wise to prove that the Markov chain would eventually converge to the posterior distribution. However, satisfying detailed balanced condition iteration-wise does not guarantee ergodicity (sampling the whole space) because of the implementation of Simulates Annealing scheme. Andrieu et al. (2000) showed Bayesian inversion using a penalized

likelihood function while using RJMCMC sampling can converge to the modes of the full posterior distribution. In our study, implementing Simulated Annealing by cooling is equivalent to using a penalized likelihood function and convergence to the modes of the posterior distribution is optimization to obtain the optimal upscaled model. The stationary distribution in our study will be a distribution of all models when the RMS error converges to the least possible value.The objective of this study is to calculate the posterior distribution $p(\mathbf{m}/\mathbf{d})$ which determines the most likely value and non-unique earth models which fit the data. If the transition kernel can choose new models m' from posterior distribution $p(\mathbf{m}/\mathbf{d})$ in such a way that $p(\mathbf{m}')K(\mathbf{m},\mathbf{m}') = p(\mathbf{m})K(\mathbf{m}',\mathbf{m})$ (detailed balance condition) is satisfied, $p(\mathbf{m}/d)$ will be a stationary distribution. The transition kernel for rjMCMC sampling is given by

$$K(\mathbf{m}',\mathbf{m}) = \alpha(\mathbf{m}',\mathbf{m})q(\mathbf{m}|\mathbf{m}')+\left[1-\int_{\mathbf{m}'\neq\mathbf{m}}\alpha(\mathbf{m}',\mathbf{m})q(\mathbf{m}|\mathbf{m}')\,\mathbf{dm}\delta_{\mathbf{m}'}(\mathbf{m})\right]. \quad (2.10)$$

The first term in the kernel distribution corresponds to probability of choosing $\mathbf{m}'$ from $\mathbf{m}$ and the second term is the probability of the random walk staying at $\mathbf{m}$. A proof of the detailed balance condition for the transition kernel of rjMCMC sampling is shown in Appendix A of Mondal et al. (2010).

2.2.4 Summary of rjMCMC upscaling

To perform rjMCMC upscaling, I define a set of moves where a move is a step during the random walk. Moves considered in this study include a stochastic perturbation of an interface boundary, a perturbation of Backus averaged value in a layer, adding a new layer, deleting a layer or a combination of some of the these steps. Algorithm 2.2.4 summarizes a pseudocode for rjMCMC based upscaling. The initial model for the random walk contains an arbitrary number of randomly

placed boundaries. Three moves namely boundary move, birth move and death move, defined in Table 2.1, are randomly chosen with equal probability during the random walk. It is during the birth and death move, layers are added and deleted by placing an additional boundary or removing a randomly chosen boundary. The RMS error, between the synthetic seismogram for the upscaled model and the reference seismogram obtained from fine layered medium, is calculated for the new model and compared with the old model. The M-H ratio shown in line 14 of Algorithm 2.2.4 is calculated and new models are accepted if $\alpha > 1$ and rejected, with a probability, if $\alpha < 1$.

2.3 Comparison of upscaling methods

2.3.1 Shallow depth interval

Our first results here consider log measurements from a 300 m interval in a test well in the North Sea to demonstrate the value of the new rjMCMC approach. The interval contains an approximately 100 m section of hydrocarbons, as well as both sharp and gradational contacts and intervals with fine layering (Figure 2.1(b)). The objective in this section is to compare existing upscaling techniques to MCMC-based upscaling techniques that are introduced in this chapter.

2.3.1.1 Results for existing Backus averaging methods

I apply conventional layer and window-based Backus averaging for V_p and density logs to the shallow depth interval as shown in Figures 2.1(a) and 2.1(b). An arbitrary number of layers with equal thickness are considered for layer-based upscaling and two different wavelength dependent intervals for window-based upscaling. 16 layers are considered here to demonstrate Layer-based Backus averaging for the shallow depth interval and the rationale behind the number of layers will be explained later in the chapter. In this study, the reference seismogram is a synthetic seismogram

Algorithm 1 Algorithm summary for Log blocking with rjMCMC sampling and VFSA

Initialize random boundary locations; compute initial RMS error $\boldsymbol{E(m_0)}$.

while $iter \leq N$ **do** ▷ N is the total number of iterations.

  Set annealing temperature: $T_{iter} = T_0 * c^{\frac{iter}{10}}$

  Draw a random number $u_{iter} \in U[0, 1]$ to select move type.

  if $u_{iter} \leq 0.33$ **then** ▷ Assigning equal probability for all three moves.

    Boundary change move

  else if $0.33 \leq u_{iter} \leq 0.66$ **then**

    Birth of layer move

  else

    Death of layer move

  end if

  Randomly select a layer and perturb the elastic properties.

  $\Delta E = E(m_{iter}) - E(m_{iter-1})$ ▷ Difference in RMS error.

  $\alpha =$ prior ratio * proposal ratio * $\exp(\frac{\Delta E}{T_{iter}})$ ▷ M-H ratio.

  Draw a random number $r \in U[0, 1]$

  if $\alpha \geq r$ **then**

    Accept m_{iter} ▷ Accept new model.

  else

    $m_{iter} = m_{iter-1}$ ▷ Reject new model.

  end if

end while

Table 2.1: Definition of moves selected during random walk and three cases allowing different combination of moves in this study.

Move	Case 1	Case 2	Case 3
Boundaries Change	Y		
Boundaries change with random perturbation of elastic properties (Boundary move)		Y	Y
Add boundary with random perturbation of elastic properties (Birth move)			Y
Delete boundary with random perturbation of elastic properties (Death move)			Y

generated in the fine layered media using propagator matrix method with source frequency of 30 Hz, assuming a Ricker wavelet.

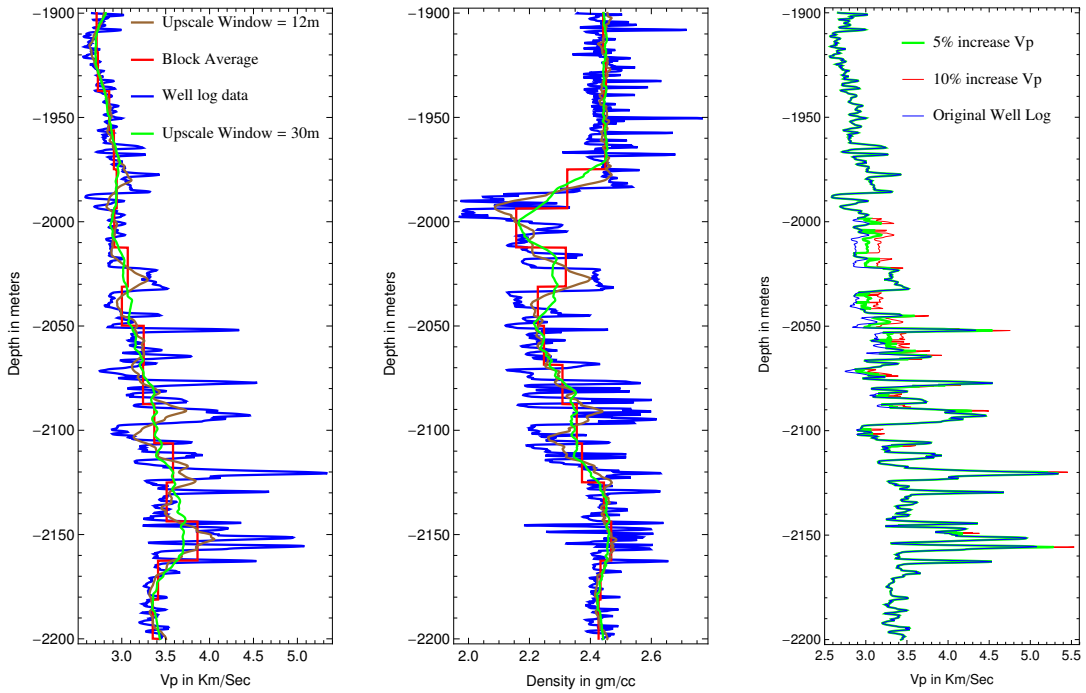
The accuracy of the upscaling techniques is measured from the difference between the upscaled model seismogram and the reference seismogram. Figure 2.2 shows the misfit between modeled and reference seismograms when using layer-based and window-based Backus upscaling. Depending on how the upscaled elastic properties are used, the misfit for a certain window length of Backus averaging can be unacceptable. For example, if the upscaled elastic properties are used in a time lapse analysis to study the change in pressure and saturation in the reservoir, the layer and window-based velocities are unacceptable as the RMS difference between base and monitor seismogram will be comparable to the RMS error due to misfit. When the V_p and density values in the oil interval are reduced by 5% and 0.5% respectively (correspond to change in pressure and Sw in time lapse applications) as shown in Figure 2.1(c), the RMS difference between the base and monitor seismogram is 5.54 . RMS difference when V_p and density in the oil interval are reduced by 10% and 1%, respectively, is 10.3. However, the RMS error between the modeled seismogram and reference seismogram calculated based on Backus-average window lengths of $\lambda/4$ (30 m) and $\lambda/12$ (12 m) are 14.49 and 3.557 respectively. The RMS error values, while upscaling, are in direct comparison with the RMS difference values for the time lapse examples which emphasizes the need for a more accurate well log upscaling method. Ramirez et al. (2013) successfully inverted for 2-D permeability field using Bayesian inference with MCMC sampling technique using synthetic data, however, the inversion with real data (i.e surface seismic and CO_2 production data) was done only with partial success. The main cause for the limited success is cited as the inaccurate elastic modeling of the geological model. The misfit between the modeled seismogram with the most accurate elastic model and observed data is larger than the deviation

of seismogram produced by the injection of CO_2.This emphasizes the need for an elastic model which can represent the geology with minimum misfit compared to the sensitivity of the seismic response with saturation or pressure changes.

Figure 2.2(d) shows how the RMS error increases with increase in the upscaling window, expressed as percentage of wavelength. Liner and Fei (2006) created a plot for image correlation error (correlation error between synthetic shot gather for fine scale well log data and Backus average equivalent medium) versus upscaling window as a percentage of minimum dominant wavelength. The correlation is approximately zero when the upscaling window is less than 1/3rd of the minimum dominant wavelength and increases rapidly when the upscaling window increases any further. A similar effect is identified in Figure 2.2(d), where the RMS error increases rapidly as the window length increases from $\lambda/10$ to $\lambda/3$ and as the upscaling window increases any further the RMS error stabilizes.

2.3.1.2 MCMC based Backus averaging methods

Layer-based upscaling results in an RMS error value of 12.9 when the interval is divided into 16 equal thickness layers. In equal thickness layers-based upscaling, the sharp contrasts of elastic properties in the middle of the layers will be artificially moved to the layer boundaries thereby resulting in incorrect travel time and amplitudes. Backus average also assumes the fine layering is isotropic and the elastic properties in the upscaling window or layer does not follow any spatial trend. Gibson and Hwang (2009), using an MCMC approach, showed an improvement in elastic modeling after allowing stochastic perturbation of each of the layer boundaries. In this study, I will show how allowing more types of moves, or model changes, can improve the predicted seismic response. Table 2.1 shows the definition of moves and three different cases which are discussed in this study. The first case is when I allow the random walk



(a) (b) (c)

Figure 2.1: (a-b) Comparison of upscaled well logs using an arbitrary number of equal layer-based (red solid) and window-based upscaling (brown and green solid). Window-based upscaling is performed with two window lengths, $\lambda/4$ (30 m) and $\lambda/12$ (12 m). (c) Original well log compared with two cases of modified log representing two separate time lapse conditions. The well log in green is after increasing the velocity in the oil interval by 5%. The increase in velocity here is assumed to correspond to a change in pressure and Sw.

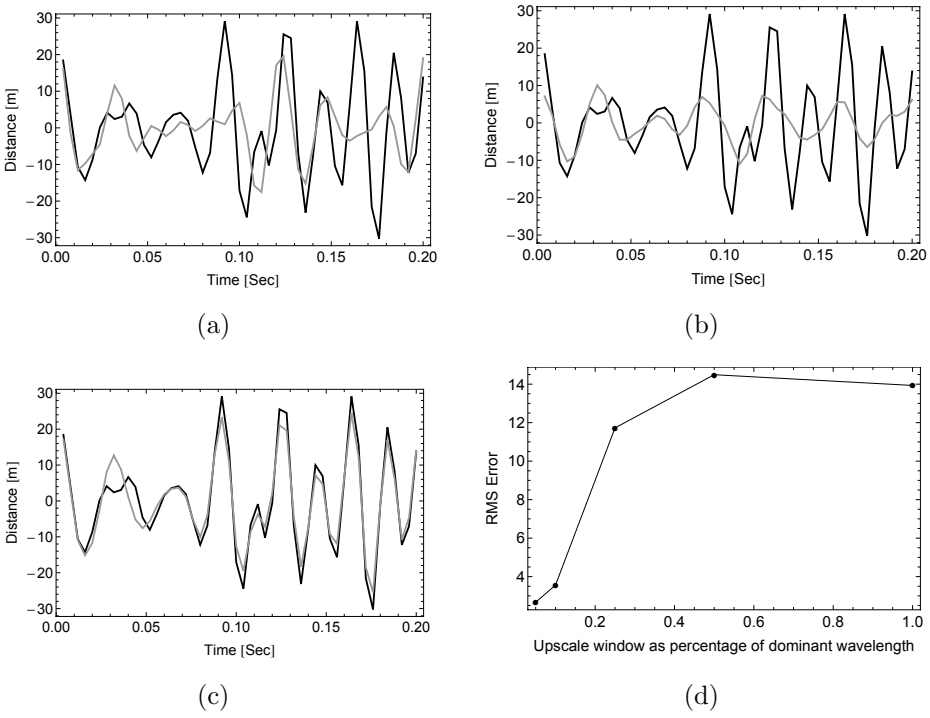


Figure 2.2: Comparison of modeled seismogram (in solid blue) and reference seismogram (in solid red) (a) Layer-based Backus average (RMS error = 12.9). (b) window based upscaling with window length = 30 m (RMS error = 14.49) (c) window based upscaling with window length = 12 m (RMS error = 3.557) (d) RMS error between reference and model seismograms versus upscaling window as a percentage of wavelength. The model seismograms are obtained based on window-based Backus averaging.

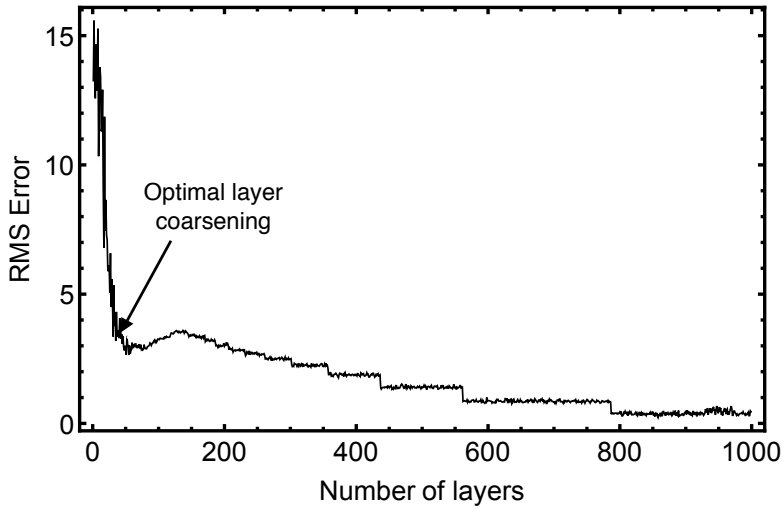


Figure 2.3: RMS error versus number of layers.

only for boundary interfaces with a fixed number of layers. The second case is when I allowed the boundary interface to change and a simultaneous random perturbation in the velocity of a layer. The third case applies three possible moves: random perturbation of boundary interface, adding new layer and deleting new layer with simultaneous perturbation of velocity during all the three steps. The third case applies a transdimensional inference due to change in dimensions permitted by adding or deleting layers during the random walk. The results from the three cases are discussed below.

2.3.1.3 Case 1: Boundary change

In the first case, a boundary is randomly chosen and allowed to change. In this case the number of layers is fixed during the inversion. One of the ambiguities while performing fixed dimensional inversion based upscaling is how to choose the number of layers. A sequential equal layer thickness Backus averaging approach can be used to gain insight about optimal number of layers for upscaling. Figure 2.3

shows a steep decrease in RMS error when the number of layer increases from 15 to 30 suggesting the optimal number of layers can be in that range. The MCMC algorithm is run for 10000 iterations using Very Fast Simulated Annealing (VFSA) approach, introduced by Ingber (1989), to assist the global optimization. The cooling schedule used in the simulated annealing process is shown in Algorithm 2.2.4. The magnitude of random movement of boundaries is reduced with iterations and is dependent on cooling schedule. The critical temperature and the cooling schedule are chosen appropriately to allow proper initial mixing and slower convergence to a global optimum and to avoid trapping of the model parameters at a local optimum. A steeper cooling schedule may accelerate the convergence to a global optimum but may result in local stagnation.

An estimate of uncertainty of the model parameters is measured from an ensemble of models. Figures 2.4(a) and 2.4(b) show the interface histograms and mean, upper and lower limits of the velocity distributions of 5000 best models. The initial model has 16 arbitrary thickness layers and the vertical velocities in layers are calculated using Backus average. The sharp contacts in the well log are identified as narrow distributions (contacts at 1985 m, 2020 m) in the interface histogram plot and the gradational contacts are wider distributions (contacts at 1970 m, 2110 m). The spread in the velocities in Case 1 is only due to the change in the layer boundaries. The velocities in layers with narrow boundary distributions have lower spread compared to that in layers with wide boundary distributions. Figure 2.5(a) shows how RMS error decreases with iterations for Case 1. Gibson and Hwang (2009) showed RMS error plots for three different data error variance values (C_d in Equation (2.2)). With increase in data error variance the uncertainty of the accepted model parameters increases and acceptance ratio will be higher. A smaller value of σ would mean that, only models that produce seismograms closer to the reference seismogram, i.e.,

smaller error, would be accepted. I chose a σ value of 2, which is approximately 3% of the peak to peak amplitude of the reference seismogram, to balance both the acceptance rate and to avoid poor fits. A depth dependent data error variance without correlation between data points can be expressed as

$$C_d = \sigma_i^2 \delta_{ij}. \tag{2.11}$$

In this study, all the diagonal elements of the data covariance matrix are equal and no spatial correlation is considered. Bodin et al. (2012) studied both exponential and Gaussian data error correlation with depth and used a Hierarchical Bayesian formulation to invert for data noise. The advantage of an exponential data error correlation function is that an analytical solution for the inverse of the data error covariance matrix exists. The RMS error value between the best model seismogram and the reference seismogram while allowing only boundary to change during the random walk is 4.55 (seismograms shown in Figure 2.6(a)) which is ∼1/3rd of the RMS error corresponding to layers of equal thickness.

2.3.1.4 Case 2: Boundary change with random elastic properties perturbation

I test the MCMC algorithm with the same well interval for Case 2, allowing the boundary and the layer velocity to change simultaneously. Similar to Case 1, the initial model has 16 arbitrary thickness layers. The perturbed layer velocities are drawn from a Gaussian distribution with Backus averaged value as mean and a standard deviation of 0.1. As observed in Figure 2.4(d), the spread in the velocity distribution is greater compared to corresponding value in Case 1 which is due to the additional layer velocity perturbation step. The RMS error after convergence for Case 2 is lower than the RMS error in Case 1 (Figure 2.6(c)). This RMS error after Case 2 is approximately 1/4th the RMS error when 16 equal thickness layers are

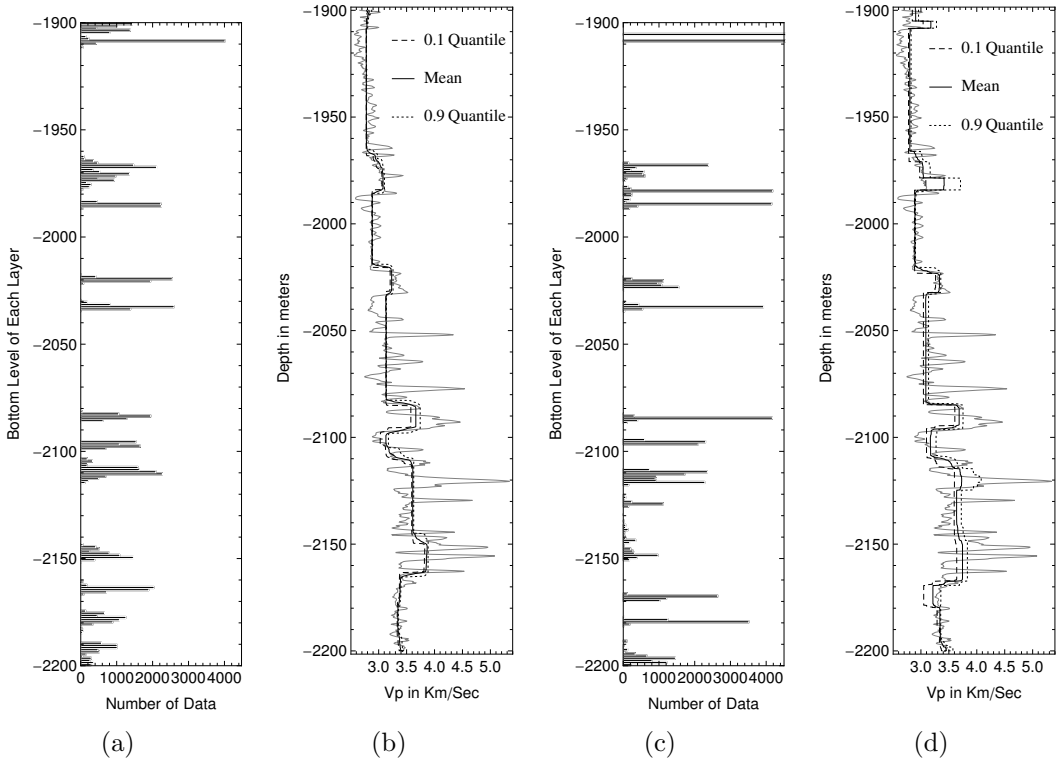
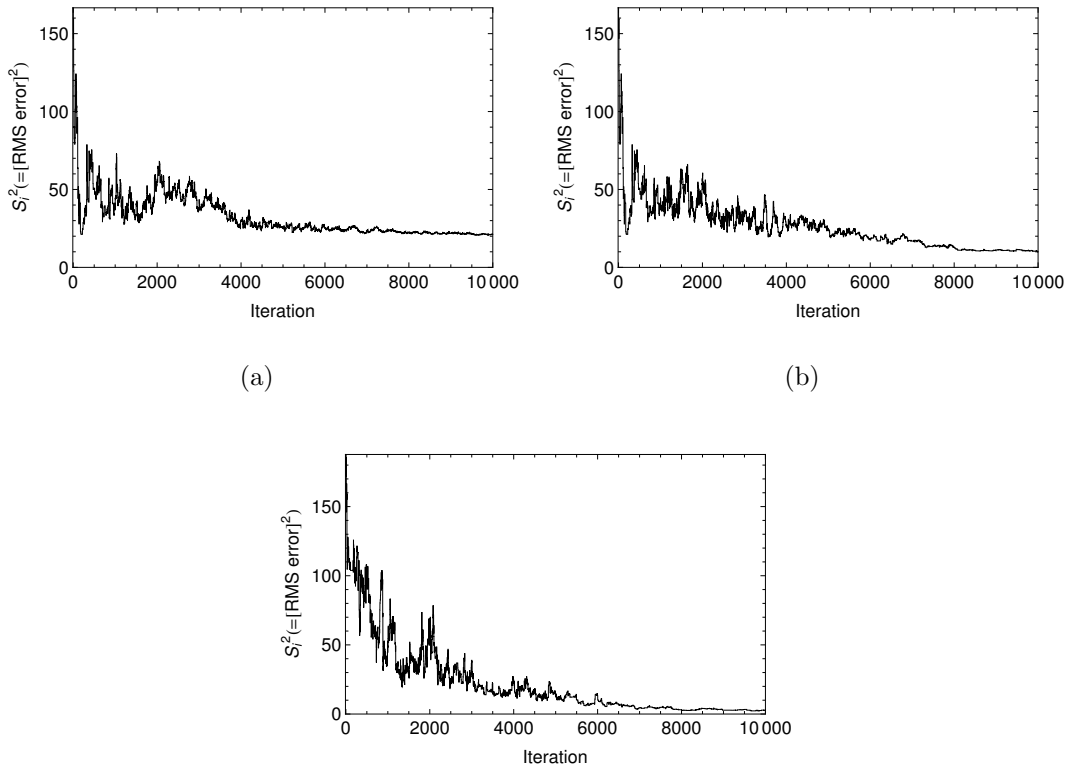


Figure 2.4: (a) Interface histograms (b) Mean, upper and lower limit of the velocity distributions after running Case 1. (c) Interface histograms (d) Mean, upper and lower limit of the velocity distributions after running Case 2. Random walk is allowed to run for 10000 iterations. The distributions are obtained from 5000 iterations with the lowest RMS error. The best model is obtained from the iteration with the least RMS Error.

considered. The layer density values are constrained to the velocity values based on a linear relationship obtained from velocity and density at fine scale from the well log. The relationship between velocity and density is $Vp = 0.1097 * rho + 2.014$ which implies a ΔV_p perturb will be accompanied by $0.1097\Delta V_p$ change in density. The number of thin layers (thickness ≤ 10 m) after Case 2, as seen in Figure 2.4(c), are higher than those in Case 1. The velocity in the thinner layers is 10-15% higher than the Backus average values compared to other layers where the deviation from Backus average is relatively low. Figure 2.7(a) shows the histogram of $\%\Delta V_p$ (Definition explained in figure caption) which suggests most of $\%\Delta V_p$ values are close to zero and there are few values near 10-20%. Figure 2.7(b) shows that all the $\%\Delta V_p$ values from 10-20% correspond to layers with thickness less than 10 m. Unrealistic layer velocities in the thinner layers can be avoided by assigning a prior for velocities and reject models with velocities higher than a prior assigned value.

2.3.1.5 Case 3: Transdimensional inference

Case 3 draws the transdimensional inference, in that the number of layers are allowed to change by allowing adding or deleting layers. Algorithm 2.2.4 shows the pseudocode used to run the transdimensional inversion for well log upscaling. There are both sharp and gradation contacts which are identified in all the three cases. The interface histograms, mean, upper and lower limit of the velocity distributions for Case 3 are shown in Figures 2.8(a) and 2.8(b). For example the gradational contact near 2110 m and sharp contacts at 2085 m and 2100 m are commonly identified in all the three cases. Figure 2.6 compares the best model seismograms, and ensemble of 5000 best model seismograms for all the three cases discussed here. The least RMS error for the best model for Case 3 is 1.48 which is approximately 1/9th of the RMS error when considering 16 arbitrary equal thickness layers. RMS error



(a)

(b)

(c)

Figure 2.5: (a) Squared error as a function of MCMC iterations for Case 1 when allowing only the boundaries to change (b) Squared error as a function of MCMC iterations for Case 2 when allowing only the boundaries to change and simultaneous random perturbation of a layer velocity (c) Squared error with all the three moves.

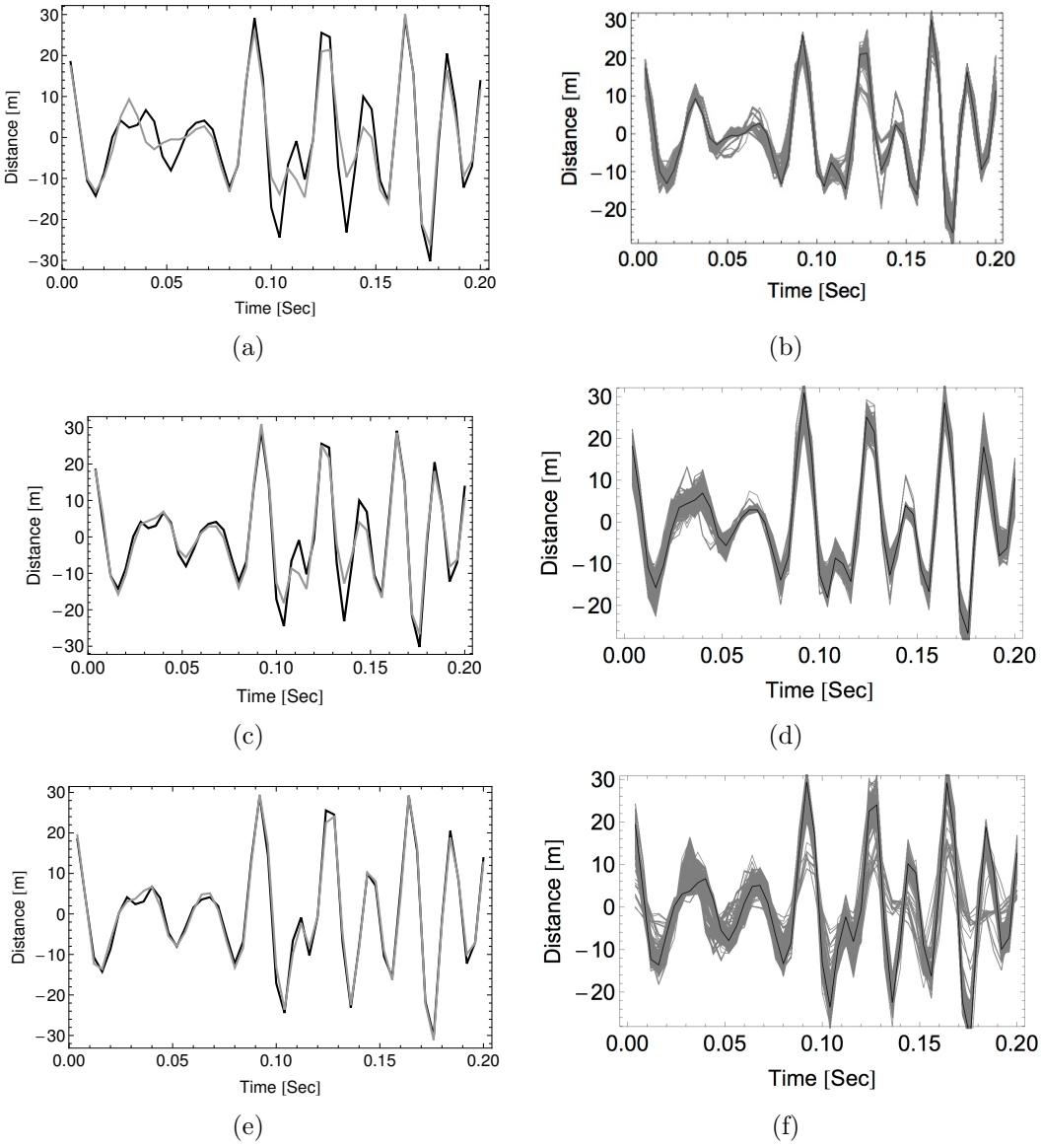
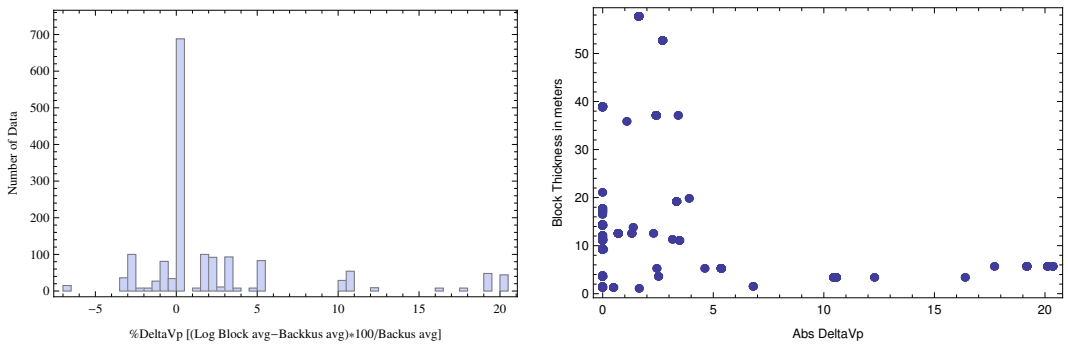


Figure 2.6: Comparison of modeled seismogram with the least RMS error (blue solid) and reference seismogram (red solid) for (a) Case 1 (RMS error = 4.55), (c) Case 2 (RMS error = 3.14) and (e) Case 3 (RMS error = 1.48). Ensemble of model seismograms for 5000 iterations with the lowest RMS error values for (b) Case 1, (d) Case 2 and (f) Case 3.



(a) (b)

Figure 2.7: (a) Histograms of $\%\Delta V_p$ after running Case 2. The $\%\Delta V_p$ value in a layer is the difference between optimal model velocity and the deterministic Backus average. The histograms include $\%\Delta V_p$ values for all the layers from the 100 best models. (b) DeltaVp versus layer thickness for all the layers in the 100 best models.

in Case 3 converges to a lower value in fewer iterations compared to Case 1 and Case 2 which is due to better mixing of boundaries introduced by the birth and delete steps. The reversible jump MCMC is run for 10000 iterations and the number of layers is allowed to change to obtain the best model with optimal number of layers. The number of layers in the random walk decreases initially and subsequently increases to higher values with a best model having 16 layers. The mixing at lower number of layers allows for low frequency inversion followed by high frequency approximation of elastic properties by increasing the number of layers and reducing the RMS error. The Bayesian formulation explained earlier is extended, and the acceptance term (Metropolis-Hastings ratio) is derived for well log upscaling, in Appendix. The acceptance term is the ratio which determines the condition for accepting/rejecting samples. The combination of prior, proposal and likelihood ratios in the acceptance term have an inherent property to chose models with fewer layers that results in a fit as good as the complex models. This property where MCMC algorithm converges to models with optimal number of parameters is called parsimony

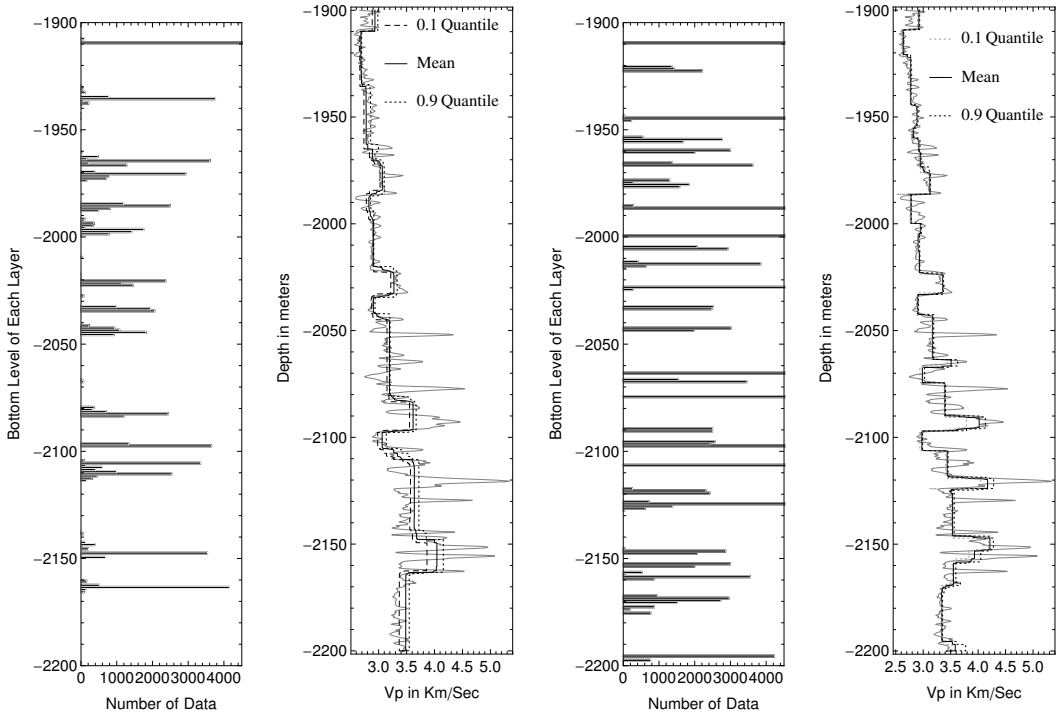


Figure 2.8: (a) Interface histograms (b) Mean, upper and lower limit of the velocity distributions after running Case 3. (c) Interface histograms (d) Mean, upper and lower limit of the velocity distributions after running Case 3 with source frequency of 60 Hz. Random walk is allowed to run for 10000 iterations. The distributions are obtained from 5000 iterations with the lowest RMS error. The best model is obtained from the iteration with the least RMS Error.

discussed by Malinverno and Leaney (2005) and Malinverno (2002). I explained how parsimony assists in sampling optimal models for our study in Appendix.

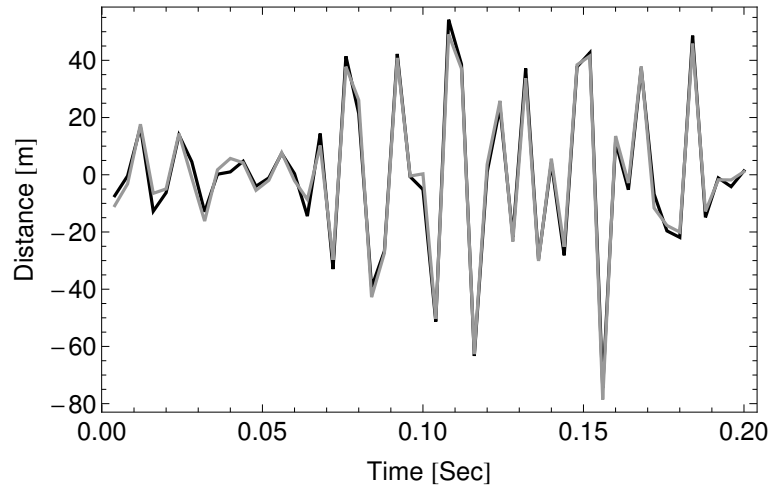
Agostinetti and Malinverno (2010) showed the optimal number of layers can be a function of data noise variance. Here, I show the number of layers and the thickness of the layers can be a function of source frequency. As the frequency increases, I would expect the optimal number of layers would increase. To verify the response of the rjMCMC sampling technique, Case 3 is run on the same 300 m interval from the test well using a source frequency of 60 Hz. Figures 2.8(c) and 2.8(d) show the

interface histograms and mean, upper and lower limits of the velocity distribution of the 5000 best models when the source frequency is 60 Hz. The number of layers of the best model is twice as much compared to the number of layers when the source frequency is 30 Hz and most of the contacts are identified as sharp contacts. The standard deviation of the velocity perturbation values and the VFSA parameters are the same as that considered for 30 Hz case. The least RMS error value for the best model seismogram is 2.91 (Figure 2.9(a)) and as seen from the RMS error plot in Figure 2.9(b), the Markov Chain converges after ∼5000 iterations. The number of layers during the random walk increases from an initial value of 10 and the best model has 32 layers. The optimal number of layers doubled when the frequency is doubled for this example. I may not always notice a direct linear relationship between frequency and the number of layers but I should expect an increase in number of layers as the frequency increases.

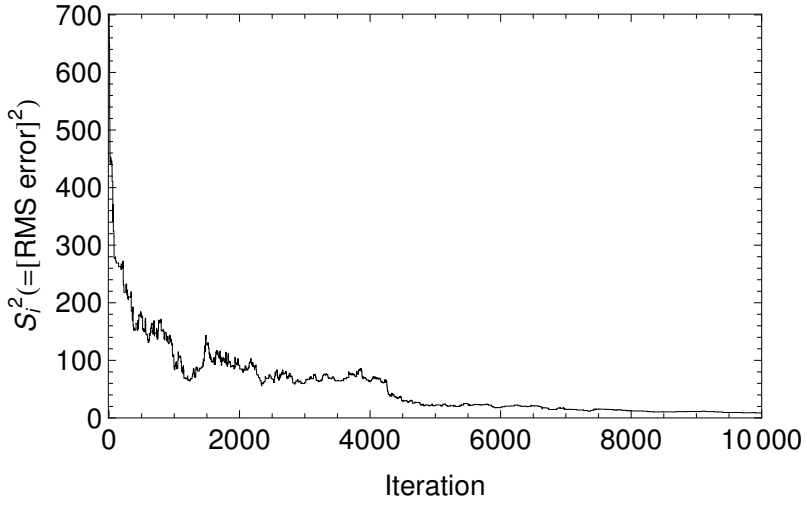
2.3.2 Deep interval

The second interval studied in this chapter is a deeper interval from 2600 m - 2740 m from the same test well in North Sea. This interval has more heterogeneities in the shallow depths compared to deep in the interval and contains hydrocarbons including thin gas beds as shown in Figure 2.10(c).

All the three cases discussed for the shallow interval and shown in Table 2.1 are also run for the deep interval but only results from case 3 are shown here. The top of the thin gas beds is identified as a sharp boundary in all the three cases at ∼2627 m and the bottom contact is similar to a Gaussian distribution in all the three cases with a mean value of distribution at ∼2645 m. There are four Gaussian like interface distributions after Case 3 at ∼ 2645 m, 2685 m, 2705 m and 2715 m respectively corresponding to bottom of the layer with thin gas beds, water oil contact, top and



(a)



(b)

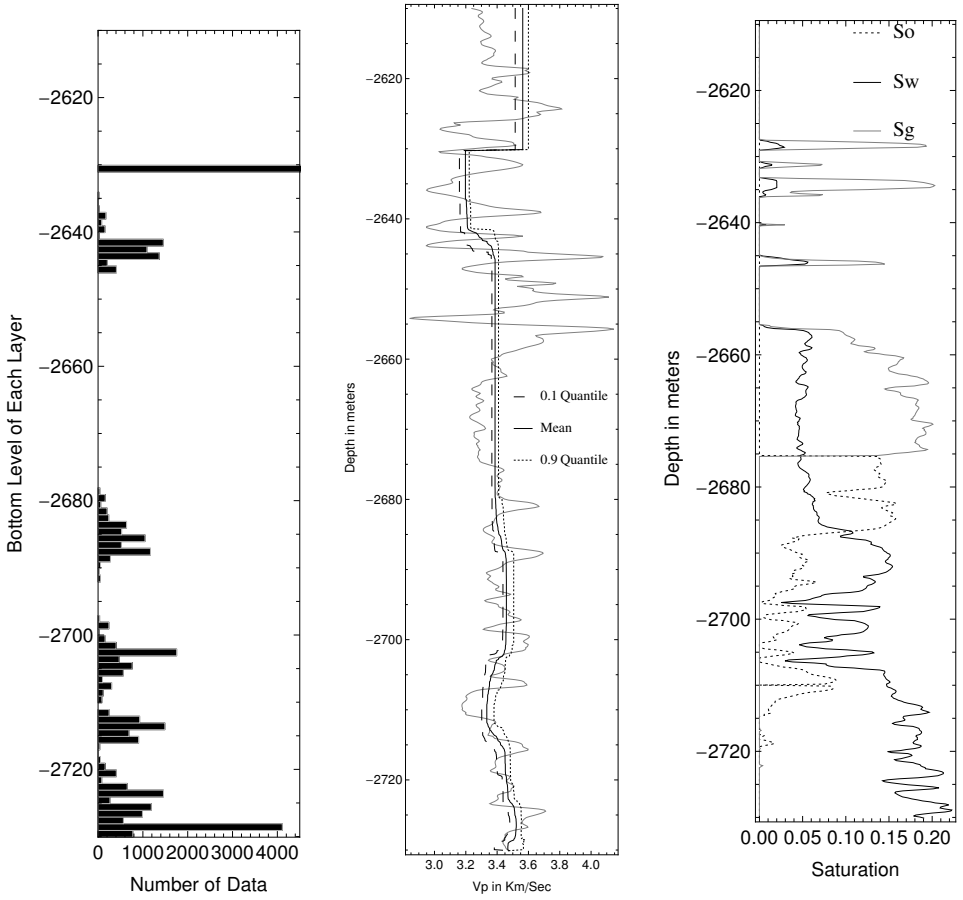
Figure 2.9: Results after running Case 3 for a source frequency of 60 Hz (a) Comparison between reference seismogram and best model seismogram (RMS error = 2.91) (b) RMS error profile with number of iterations.

bottom of interval with thins beds of oil (Figure 2.10). In the deeper interval, the criteria for layer boundaries mainly seems to be litho-fluid classes compared to the shallow interval where they are chosen based on lithology. An important outcome from transdimensional inference, based on the results from both the intervals, is rjMCMC sampling determines the optimal number of parameters, to fit the data, and also ensures the boundaries are placed based on geologic evidence. Similar to the analysis of shallow interval, the number of layers with iterations initially decreases from starting value of 10 layers and the best model has an optimal number of 8 layers.

2.4 Discussion

The distributions of layer boundaries based on the samples collected after burn-in resemble a sharp, Gaussian or a uniform distribution. A sharp distribution implies a sharp contrast in the velocity log representing a sudden change in lithology. A gradational contact is usually represented by a Gaussian distribution. A uniform distribution suggests the uncertainty in the layer boundary is large and the presence of the layer boundary probably does not effect the model seismogram. Uniform distributions are identified at $\sim$1900 m and $\sim$2200 m as shown in Figure 2.4(a) and near 2200 m in Figure 2.4(c) where these results are obtained without allowing the number of layers to change. However, in the case when the number of of layers are allowed to change (Figure 2.8(a)) the layer boundaries corresponding to the uniform distributions disappear.

De et al. (1994) compared log velocities with VSP velocities and found VSP velocities are consistently lower than high frequency velocities and attributed this difference to velocity dispersion. In this study, the deviation of layer velocities from Backus average values varies from -5% to 8% based on values obtained from both shallow and deep intervals as shown in the histograms of Figures 2.11(a) and 2.11(b).



(a) (b) (c)

Figure 2.10: (a) Interface histograms (b) Mean, upper and lower limit of the velocity distributions after running Case 3 for the deeper interval. (c) Oil, gas and water saturation calculated for deep interval. Random walk is allowed to run for 10000 iterations. The distributions are obtained from 5000 iterations with the lowest RMS error. The best model is obtained from the iteration with the least RMS Error.

$\%\Delta V_p$ values are both positive and negative and most of the values are close to zero as observed from the histograms. Unlike the $\%\Delta V_p$ histogram data by De et al. (1994), Goetz et al. (1979) showed $\%\Delta V_p$ values are both positive and negative with mostly negative values for shallower depths (<3000 ft) and negative for deeper intervals (>3000 ft). Our results concur with observations by Goetz et al. (1979) but no positive or negative trend with depth are identified. De et al. (1994) used an arithmetic smooth window averaging, not Backus averaging, to upscale well logs; and arithmetic averaging, like Voigt averaging, result in the upper limit of effective medium parameters and this might be partially responsible for the consistent positive deviation of high frequency velocities from VSP derived velocities. After obtaining the $\%\Delta V_p$ values, I are motivated to investigate if they are dependent on any wave propagation or geologic parameters. I cross plotted $\%\Delta V_p$ values with various elastic and petrophysical parameters.

The layer velocities with higher hydrocarbon saturations are closer to the Backus average values as observed in Figure 2.11(c) and I observed similar trend for plots obtained from the deeper interval. To examine the effect of rapid and gradual variations (i.e proxies to depositional environments of the sediments) in the velocity of the well logs on upscaling I compared the standard deviation of V_p values in the layers and $\%\Delta V_p$ (Figure 2.11(d)) and no correlation is observed. I made similar plots for different vertical spatial correlations (not shown here) versus $\%\Delta V_p$ and no relationship is identified.

2.5 Conclusions

Well log upscaling using transdimensional inversion results in the simplest equivalent medium model which can propagate the wave similar to the fine layered medium. Earlier literature on Backus averaging for well log upscaling suggest only the present

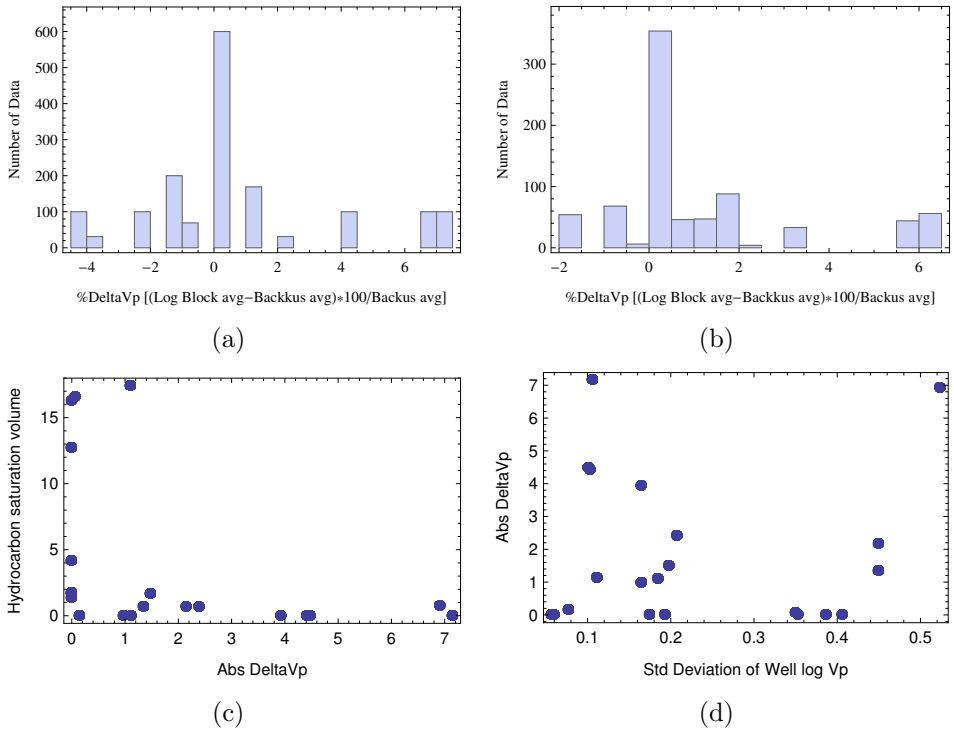


Figure 2.11: (a) Interface histograms (b) Mean, upper and lower limit of the velocity distributions after running Case 3 for the deeper interval. (c) Oil, gas and water saturation calculated for deep interval. Random walk is allowed to run for 10000 iterations. The distributions are obtained from 5000 iterations with the lowest RMS error. The best model is obtained from the iteration with the least RMS Error.

study and numerical modeling by Grechka (2003) quantify deviation of optimal velocities of the equivalent medium from Backus average. The approach in this chapter removes the ambiguity of choosing an upscaling window by random movement of the interface and also allowing the number of layers to change. The best model interfaces correspond to contrasts in lithology or litho-fluid classes. Sharp contrasts in the velocity log are identified as sharp distributions of boundaries in the random walk and gradational contacts are identified as wide distributions.

3. QUANTIFICATION OF UNCERTAINTY IN VELOCITY LOG UPSCALING USING REVERSIBLE JUMP MCMC

3.1 Introduction

Full waveform inversion (FWI) improves the confidence of the earth models obtained during inversion using seismic data compared to the models obtained while using only traveltime as reference data. FWI aims to reduce the mismatch between the model seismograms generated during the random sweep of model parameters and a reference seismogram. FWI can be defined in a stochastic framework, where velocity profile at an iteration is referred as a model. Stochastic FWI is a multi-parameter inversion technique where the objective is to find a global minimum in the model space.

One of the widely used applications of FWI is to generate velocity of earth models for seismic imaging using normal incidence or pre-stack seismic data as reference data. The velocity inversion procedure involves using a rough initial guess followed by iteratively perturbing the velocity model to obtain an accurate representation of the earth model. Often, well logs can be used to constrain and stabilize the inversion results. Upscaled velocity log values can provide a good estimate for an initial guess velocity model. This chapter addresses upscaling velocity logs as a waveform-based inversion problem by comparing the seismic signal from the upscaled models with that from fine layered medium. The upscaled velocity in each layer is a perturbation of velocity from the deterministic Backus average value. Gibson and Hwang (2009) referred to this approach of upscaling as log blocking because of the block-based nature of upscaling where block in this case is a layer. Since the inversion is performed in a Bayesian inversion framework, the results also include uncertainty in the upscaled

velocity models.

In this chapter, I compare velocity log upscaling results when using reversible jump MCMC with and without Simulated Annealing (SA) to guide optimization. The results obtained when using SA (discussed in chapter 2) are obtained by tempering the likelihood function which allows the sampler to rapidly converge to the modes of the full posterior distribution and give no description of the distribution. Although using SA can assist in rapid sampling of optimal models, the sampler can also get trapped in a suboptimal local minimum. Reversible jump MCMC without using SA samples models from the full posterior distribution, but estimating the optimal model requires sampling many more models, and hence more forward model evaluation, are needed. I compare the uncertainty in upscaled velocities for two well log intervals from a well in North Sea with and without tempering the likelihood function in the Bayesian inversion framework. I extended log blocking, to obtain upscaled velocities using multi-offset seismograms as reference seismograms, to a well log measured at location U1347 on the Shatsky Rise oceanic plateau. I introduced an empirical convergence assessment for rjMCMC sampler to determine a point in the random sweep where the sampler had ‘‘converged”. Convergence here refers to the point after which the samples are generated from the full posterior distribution.

Below I summarize log blocking while using Bayesian inversion with and without tempered likelihood function and compare the results from both the approaches. I also summarize the results of log blocking when using Multi-offset seismogram as a reference seismogram. Finally, I discussed a univariate and multivariate empirical diagnostic to check convergence while using rjMCMC sampling.

3.2 Methods

In this chapter, I compare results from two approaches of log blocking in a Bayesian inversion framework. The two approaches include log blocking with and without a tempered likelihood function. Log blocking is a layer-based upscaling technique where upscaling is defined as a Bayesian inversion problem. All the results in this chapter are obtained by implementing rjMCMC sampling which allows the number of layers to change during the random sweep and thereby making the inversion a transdimensional inverse problem. The layer boundaries and layer velocities are additional model parameters which are allowed to change during log blocking and a detailed procedure on moves during rjMCMC sampling are discussed in chapter 2. The upscaled layer velocities are calculated by using Backus averaging of fine layered velocities (Velocity log) followed by a random perturbation. In a Bayesian inversion framework, the prior information is updated based on a likelihood function to derive a posterior distribution of the model parameters. The model parameter vector $\mathbf{m}$ includes $(k, \mathbf{z}, \mathbf{v})$ with their definitions same as in chapter 2. The prior for k is a uniform distribution given by $1/(k_{max} - k_{min})$ with a lower (k_{min}) and higher (k_{max}) limit. The prior for $\mathbf{z} = (z_1, ...z_{k-1})$ is given by

$$p(\mathbf{z}) = \prod_{i=1}^{k} \omega/(z_{i+1} - z_{i-1}), \tag{3.1}$$

where ω indicates coefficient of step size which defines the amount of layer boundary perturbation during the random sweep that must be chosen by testing multiple values. The perturbation of a layer boundary depends on the adjacent layer thickness values. Larger thickness values of adjacent layers implies larger perturbation values of the layer boundary. The prior for number of layers is chosen as a non-informative prior

to let the data guide the inversion. A uniform prior distribution for layer velocities is given by $\frac{1}{\Delta v_i}$, where i is the layer number allowed to perturb and $\Delta v_i = v_{max} - v_{min}$.

3.2.1 Log blocking with tempered likelihood function

The likelihood function is a way to measure the mismatch between the reference and model seismogram. Simulated Annealing is a physical process in which initial (critical) temperate is chosen to allow the sampler to traverse a large volume of the model space and when the temperature reduces gradually (slow cooling) precise sampling near global minimum is possible. Slow cooling to assist the sampler in finding the global minimum can be implemented by introducing a temperature term in the likelihood function. This procedure of including temperature in the likelihood function is referred to here as tempering a likelihood function. The temperature function for slow cooling in this study is given by

$$T_{iter} = T_0 * c^{\frac{iter}{10}}. \tag{3.2}$$

where T_0 is the initial temperature, c is the cooling constant and $iter$ is the iteration number during the random sweep. T_0 should be chosen appropriately to allow the initial samples to be well dispersed in the model space. c is usually a fraction value close to 1 and defines the rate of cooling. The tempered likelihood function is

$$P(\mathbf{d}|\mathbf{m}) = \frac{1}{(2\pi)^{(n/2)}|\mathbf{C_d}|^{(1/2)}} \exp\left[\frac{-\frac{1}{2}(g(\mathbf{m}) - \mathbf{d})^T \mathbf{C_d}^{-1}(g(\mathbf{m}) - \mathbf{d})}{T_{iter}}\right]. \tag{3.3}$$

where $\mathbf{C_d}$ is the data covariance error matrix, and $\mathbf{g(m)}$ is the forward model function representing the synthetic seismogram for upscaled model $\mathbf{m}$. The term in the exponential of the likelihood function without the temperature term is the the error function or mismatch term ($\mathrm{E}(\mathbf{m_i})$) measured as RMS error between the

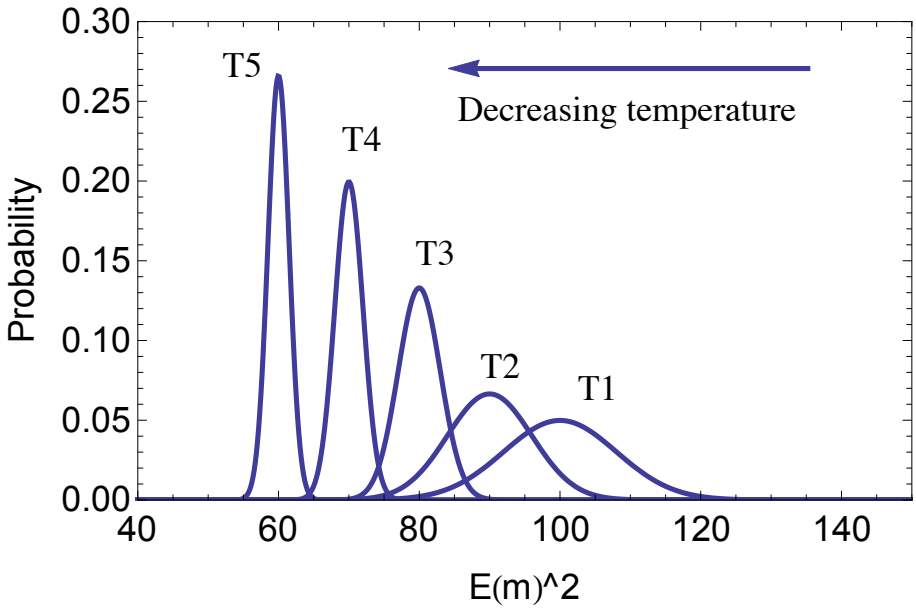


Figure 3.1: Schematic showing tempered likelihood function versus mismatch between reference and model seismogram. The mismatch reduces as the temperatures reduces with iterations.

reference seismogram and the model seismogram at each iteration. In the Markov Chain, models are accepted or rejected based on difference in mismatch between new and old models, ΔE $(E(\mathbf{m_{i+1}}) - E(\mathbf{m_i}))$. When $\Delta E < 0$ the new model is accepted and when $\Delta E > 0$ the new model is accepted with a probability. A detailed derivation of the acceptance criteria when using reversible jump MCMC sampling is discussed in the Appendix. Figure 3.1 shows how the mismatch between reference and model seismograms reduces with increase in temperature. The temperature term is derived from the cooling function as shown in equation 3.2 which decreases with iterations at a rate dependent on the cooling rate. Another observation from Figure 3.1 is how the probability distribution of mismatch changes with temperature. Initially when the temperature is high, a wide distribution for mismatch suggests well dispersed upscaled models sampled from a large volume of model space are more

often accepted compared to the case when the the mismatch distribution is narrow after cooling. A narrow distribution for mismatch implies most of the samples in the model space are rejected, however, the accepted samples will be generated at the modes of the posterior distribution and thereby generate the optimal models for upscaled medium. Andrieu et al. (2000) proved theoretically and empirically that rjMCMC sampler when used with Simulated Annealing i.e. a tempered likelihood function can converge to modes of the posterior distribution.

3.2.2 Log blocking without tempered likelihood function

The methodology of log blocking without tempered likelihood function is similar to the earlier method except temperature cooling is avoided. The likelihood function in this case is

$$P(\mathbf{d}|\mathbf{m}) = \frac{1}{(2\pi)^{(n/2)}|\mathbf{C_d}|^{(1/2)}} \exp\left[-\frac{1}{2}(g(\mathbf{m})-\mathbf{d})^T {\mathbf{C_d}}^{-1}(g(\mathbf{m})-\mathbf{d})\right]. \tag{3.4}$$

where all the parameters mean the same as in log blocking with tempered likelihood function. The algorithm in this case does not force the sampler to accept samples only in limited model space instead it allows ergodicity (sampling the whole space) and estimates the true uncertainty of the model parameters.

3.2.3 RJMCMC empirical convergence assessment

Assessing convergence of a MCMC sampler is a crucial step to determine a point in the Markov chain when the models are sampled from the posterior distribution. Traditionally, it is a common practice to burn-in samples after the random sweep and use the remaining samples to draw inference. The objective of this section is to introduce an empirical convergence diagnostic when using rjMCMC sampling which can assist in selecting a point in the Markov chain whereafter the chain has

"converged". Converged in this context refers to the point after which the samples are drawn from the full posterior distribution.

The convergence diagnostic discussed here is only valid to determine convergence when using rjMCMC sampling without SA i.e, when likelihood function does not change with iterations. Since the objective of rjMCMC sampling with SA in this dissertation is to determine the optimal upscaled model and not to infer information about the full posterior distribution, the convergence diagnostic mentioned here is irrelevant. Further discussion of convergence in this section is valid only for rjMCMC without SA.

During MCMC sampling, effective sampling of the parameter space is verified by calculating an autocorrelation function (ACF) from samples obtained in a Markov chain. If the autocorrelation function remains high for large lag, then the sampler is sampling from a limited model space and the effective sample size to draw inference about the posterior will be low. On the contrary, if the ACF drops close to zero for small lag, then the samples are well dispersed and the effective sample size will be higher. This approach of using ACF to derive inference about effective sample size is not valid for rjMCMC sampling when the dimension of the model space varies with sweeps. A detailed description about the calculation of ACF and its uses are discussed in chapter 11 of Hoff (2009).

A simple way to verify convergence for rjMCMC sampling is by checking frequency of model dimension visits in a pooled-chain. In this approach Markov chains are run in multiple chains with well dispersed initial seeds. Well dispersed is equivalent to randomly choosing different number of layers and layer boundary locations for the initial seed models. If there is a significant difference in frequency of number of layers visited in the multiple chains, this condition implies all the samples are not drawn from the same stationary distribution.

Castelloe and Zimmerman (2002) introduced a rigorous pool-chain convergence diagnostic where multiple Markov chains are run and the variance of pooled-chain estimates is compared with variance of individual chains. If the variance estimates from pooled-chains and individual chains have similar values then all the Markov chains are sampling from the same posterior distribution. Castelloe and Zimmerman (2002) extended the convergence diagnostic introduced by Gelman and Rubin (1992) from univariate analysis to multivariate (multiple model parameters monitored). Before I discuss the pooled-chain and within-chain variance parameters, I introduce some of the sub parameters which are important to understand the variance estimates.

$$R_{\cdot m} = \sum_{c=1}^{C} R_{cm}, \tag{3.5}$$

$$\bar{\boldsymbol{\theta}}^{\cdot}_{cm} = \frac{1}{R_{cm}} \sum_{r=1}^{R_{cm}} \boldsymbol{\theta}^{r}_{cm}, \tag{3.6}$$

$$\bar{\boldsymbol{\theta}}^{\cdot}_{c\cdot} = \frac{1}{T} \sum_{m=1}^{M} \sum_{r=1}^{R_{cm}} \boldsymbol{\theta}^{r}_{cm}, \tag{3.7}$$

$$\bar{\boldsymbol{\theta}}^{\cdot}_{\cdot m} = \frac{1}{R_{\cdot m}} \sum_{c=1}^{C} \sum_{r=1}^{R_{cm}} \boldsymbol{\theta}^{r}_{cm} \text{ and} \tag{3.8}$$

$$\bar{\boldsymbol{\theta}}^{\cdot}_{\cdot\cdot} = \frac{1}{CT} \sum_{c=1}^{C} \sum_{m=1}^{M} \sum_{r=1}^{R_{cm}} \boldsymbol{\theta}^{r}_{cm}, \tag{3.9}$$

where, $\boldsymbol{\theta}$ is the vector of velocities at arbitrary depths but remained as the same across all the models; C is the number of chains; T is the number of random sweeps in a chain used to calculate the convergence parameters; M is the number of models visited with a particular number of layers in any chain; $\boldsymbol{\theta}^{r}_{cm}$ is the velocity vector in

the r^{th} occurrence of a model with m layers in a chain c. R_{cm} is the number of times a model with m layers is visited in a chain c. $\bar{\boldsymbol{\theta}}^{\cdot}_{cm}$ is a vector where each element is an average of $\boldsymbol{\theta}$ values in all models with m layers in a chain c. $\bar{\boldsymbol{\theta}}^{\cdot}_{c.}$ is a vector with C elements with each element being the average of $\boldsymbol{\theta}$ values in all the models in any chain. $\bar{\boldsymbol{\theta}}^{\cdot}_{.m}$ is a vector with M elements with each element being the average of $\boldsymbol{\theta}$ values of all the models with m number of layers. $\bar{\boldsymbol{\theta}}^{\cdot}_{..}$ is the average of $\boldsymbol{\theta}$ values of all the models across all the chains.

3.2.3.1 Univariate version

$$\hat{V}(\theta) = \frac{1}{CT-1}\sum_{c=1}^{C}\sum_{m=1}^{M}\sum_{r=1}^{R_{cm}}(\theta^{r}_{cm} - \bar{\theta}^{\cdot}_{..})^2, \tag{3.10}$$

$$W_c(\theta) = \frac{1}{C(T-1)}\sum_{c=1}^{C}\sum_{m=1}^{M}\sum_{r=1}^{R_{cm}}(\theta^{r}_{cm} - \bar{\theta}^{\cdot}_{c.})^2, \tag{3.11}$$

$$W_m(\theta) = \frac{1}{CT-M}\sum_{c=1}^{C}\sum_{m=1}^{M}\sum_{r=1}^{R_{cm}}(\theta^{r}_{cm} - \bar{\theta}^{\cdot}_{.m})^2 \text{ and} \tag{3.12}$$

$$W_mW_c(\theta) = \frac{1}{C(T-M)}\sum_{c=1}^{C}\sum_{m=1}^{M}\sum_{r=1}^{R_{cm}}(\theta^{r}_{cm} - \bar{\theta}^{\cdot}_{cm})^2, \tag{3.13}$$

where θ is a scalar velocity at arbitrary depth which is same across all models. $\hat{V}(\theta)$ can be interpreted as variance of samples in all the chains, $W_c(\theta)$ is the average of sample variance within each chain implying it is a measure of within chain variance, $W_m(\theta)$ is the variance of samples within models with same number of layers from all the chains. $W_m(\theta)W_c(\theta)$ is the variance of models with same number of layers within each chain. If the design of the Markov chain is appropriate, $\hat{V}(\theta)$ and $W_c(\theta)$ should approximate the true variance of the posterior distribution for θ and their

ratio should be close to 1. Brooks and Gelman (1998) mentioned a cutoff value of 1.2 for the ratio in the results derived in their work. Similarly, $W_m(\theta)$ and $W_m(\theta)W_c(\theta)$ should approach the true variation of samples within models with same number of layers. A detailed description and derivations of the variance parameters are discussed in Castelloe and Zimmerman (2002).

3.2.3.2 Multivariate version

$$\hat{V}(\boldsymbol{\theta}) = \frac{1}{CT-1}\sum_{c=1}^{C}\sum_{m=1}^{M}\sum_{r=1}^{R_{cm}}(\boldsymbol{\theta}^r_{cm} - \bar{\boldsymbol{\theta}}^{\cdot}_{\cdot\cdot})(\boldsymbol{\theta}^r_{cm} - \bar{\boldsymbol{\theta}}^{\cdot}_{\cdot\cdot})', \tag{3.14}$$

$$W_c(\boldsymbol{\theta}) = \frac{1}{C(T-1)}\sum_{c=1}^{C}\sum_{m=1}^{M}\sum_{r=1}^{R_{cm}}(\boldsymbol{\theta}^r_{cm} - \bar{\boldsymbol{\theta}}^{\cdot}_{c\cdot})(\boldsymbol{\theta}^r_{cm} - \bar{\boldsymbol{\theta}}^{\cdot}_{c\cdot})', \tag{3.15}$$

$$W_m(\boldsymbol{\theta}) = \frac{1}{CT-M}\sum_{c=1}^{C}\sum_{m=1}^{M}\sum_{r=1}^{R_{cm}}(\boldsymbol{\theta}^r_{cm} - \bar{\boldsymbol{\theta}}^{\cdot}_{\cdot m})(\boldsymbol{\theta}^r_{cm} - \bar{\boldsymbol{\theta}}^{\cdot}_{\cdot m})' \text{ and} \tag{3.16}$$

$$W_mW_c(\boldsymbol{\theta}) = \frac{1}{C(T-M)}\sum_{c=1}^{C}\sum_{m=1}^{M}\sum_{r=1}^{R_{cm}}(\boldsymbol{\theta}^r_{cm} - \bar{\boldsymbol{\theta}}^{\cdot}_{cm})(\boldsymbol{\theta}^r_{cm} - \bar{\boldsymbol{\theta}}^{\cdot}_{cm})'. \tag{3.17}$$

The definitions of $\boldsymbol{\theta}$, $\hat{V}(\boldsymbol{\theta})$, $W_c(\boldsymbol{\theta})$, $W_m(\boldsymbol{\theta})$ and $W_m(\boldsymbol{\theta})W_c(\boldsymbol{\theta})$ are same as mentioned earlier. However, all of them are vectors in this case as multiple parameters are monitored using the multi version version of convergence assessment.

3.2.3.3 Convergence diagnostic parameters

Brooks and Gelman (1998) and Castelloe and Zimmerman (2002) defined the ratios $\hat{V}(\theta)/W_c(\theta)$ and $W_m(\theta)/W_m(\theta)W_c(\theta)$ as potential scale reduction factors (PSRF's) given by,

$$PSRF_1(\theta_i) = \frac{\hat{V}(\theta_i)}{W_c(\theta_i)}, \tag{3.18}$$

$$PSRF_2(\theta_i) = \frac{W_m(\theta_i)}{W_mW_c(\theta_i)}, \tag{3.19}$$

$$MPSRF_1(\boldsymbol{\theta}) = \text{maximum eigenvalue of } [W_c(\boldsymbol{\theta})]^{-1}\hat{V}(\boldsymbol{\theta}) \text{ and} \tag{3.20}$$

$$MPSRF_2(\boldsymbol{\theta}) = \text{maximum eigenvalue of } [W_m W_c(\boldsymbol{\theta})]^{-1} W_m(\boldsymbol{\theta}), \tag{3.21}$$

Summary of steps followed during convergence assessment :

1. Identify a velocity profile at a random sweep which will be referred to as a model hereafter. Select $\boldsymbol{\theta} = (V_{p_1}, ..., V_{p_n})^{'}$, which will be a vector of velocities at arbitrary number of depths for this study or θ in case of a univariate version of convergence assessment which will be a scalar velocity value at an arbitrary depth.

2. Run rjMCMC sampling for log blocking in C > 1 chains with dispersed initial seeds. Dispersed initial seeds for this study will be initial models with arbitrary number of layers and layer boundaries for all the initial seeds. The upscaled velocities in each layer of the initial seed model will calculated using deterministic Backus averaging of the fine scale log measurements.

3. Choose a batch size q preferably a factor of T (for example $q \approx \frac{T}{20}$). Calculate the convergence parameters $MPSRF_1$ and $MPSRF_2$ for the samples

$$(\boldsymbol{\theta}_1^1, ...\boldsymbol{\theta}_1^{nq}), ..., (\boldsymbol{\theta}_C^1, ...\boldsymbol{\theta}_C^{nq})$$

where $n = 1....\frac{T}{q}$ is the batch number and C is the number of chains.

4. Plot the following to assess the convergence

   1. maximum eigenvalues of $MPSRF_1$ and $MPSRF_2$ versus batch number (n).

   2. maximum eigenvalues of $\hat{V}(\boldsymbol{\theta})$ and $W_c(\boldsymbol{\theta})$ together versus batch number (n).

   3. maximum eigenvalues of $W_m(\boldsymbol{\theta})$ and $W_m W_c(\boldsymbol{\theta})$ together versus batch number (n).

In the case of univariate assessment, the plots will be $PSRF_1$, $PSRF_2$, $\hat{V}(\theta)$, $W_c(\theta)$, $W_m(\theta)$ and $W_m W_c(\theta)$ versus n.

5. Identify a point in the random sweep when $MPSRF_1$ and $MPSRF_2$ reach closer to one and plots in steps 4.2 and 4.3 would approximately reach common values.

6. Collect all the samples after the point of convergence and draw inference using these samples.

3.3 Comparison of upscaling results with and without SA

In this section, I apply log blocking with and without tempered likelihood function to log measurements from a 300 and 120 m intervals from a test well in North Sea. Both the intervals have sections of hydrocarbon intervals, interbedded with thin gas layers and various sharp and gradual contacts as shown in Figure 3.2. The 300 m interval is the shallow interval and the 120 m interval is the deep interval as evident from the depths shown for the intervals. The objective in this section is to compare upscaled velocities obtained from both the methods of log blocking discussed in this chapter. Figure 3.2 shows the initial upscaled velocities of both the intervals. The upscaled layer velocities of the initial model are calculated using Backus averaging and the initial upscaled velocity model has 10 layers in both the intervals. The prior distribution of velocity is a uniform distribution with upper and lower limits as linear velocity-depth trend in the log intervals $\pm\, \Delta V_p$ where ΔV_p is half the range of the prior uniform distribution of velocity. The same initial velocity models are used for log blocking with and without tempered likelihood function.

The reference seismogram is a synthetic seismogram generated in the fine layered media using propagator matrix (Gibson Jr, 2005) with a source frequency of 30 Hz, assuming a Ricker wavelet. Figure 3.3 shows the reference seismogram and the initial model seismogram obtained from the initial upscaled velocity model for both the

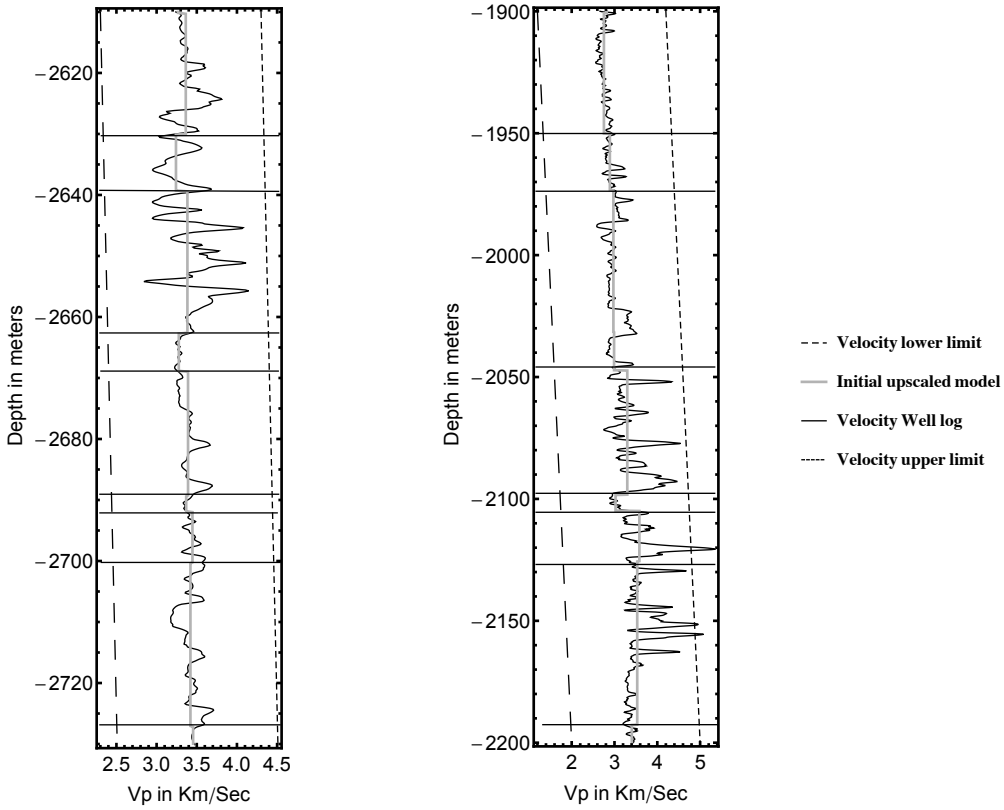
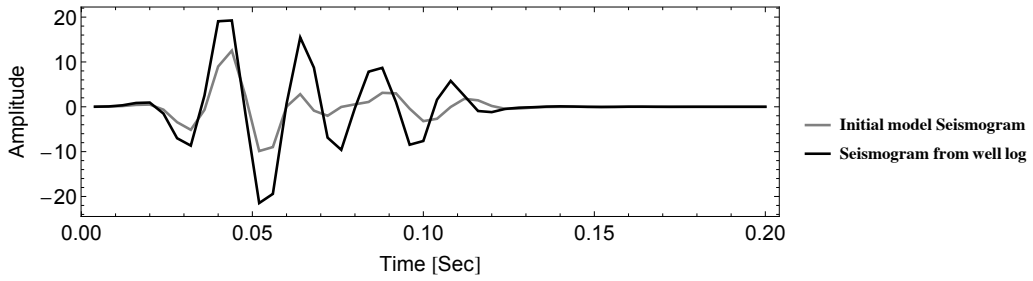


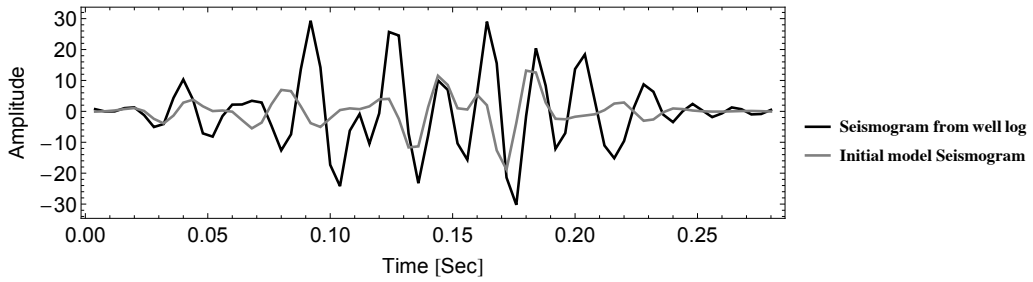
Figure 3.2: Upscaled velocity profile of the initial seed for (a) deep and (b) shallow interval. The upper and lower limit of the uniform prior distribution of velocities are also shown.

intervals. The stochastic parameters like standard deviation of the proposal function and coefficient of step size, the error standard deviation σ are remained the same while comparing results from log blocking with and without tempered likelihood function. However, these parameters may change for each interval, for example, the error standard deviation is chosen based on the maximum amplitude or energy of the reference seismogram. The only parameter different while running log blocking using both the methods discussed here is cooling temperature which is included only while using a tempered likelihood function. The standard deviation of the proposal distribution defines the magnitude of jump in the model space from the current model and the coefficient of step size defines the magnitude of layer boundary perturbation during the random sweep.

Figures 3.4 and 3.6 show how the mismatch between reference and model seismogram varies with iterations for both the intervals during log blocking. In the case of log blocking with SA the mismatch in first 5000 iterations is similar to the case of log blocking without SA, however, subsequently the mismatch reduces to a value less than 3 for both the intervals. In the first 5000 iterations, due to the selected cooling rate, the temperature is allowing the sampler to accept samples from a large volume of model space and eventually as the temperature further decreases, the sampler is accepting samples only near the modes of the posterior distribution. Figure 3.5 shows the ensemble of upscaled models in a running window of 1000 iterations at a thinning interval of 200 iterations. There is a low frequency trend in the RMS error function as observed in Figures 3.4 and high frequency variation in the RMS error around the low frequency trend. Comparing the upscaled models in the first 3000 iterations with upscaled models after 3000 iterations, the drop in RMS error at ~ 3000 iterations is due to the increase in the number of layers. The high frequency fluctuations is due to the perturbation of boundaries. For example, when one or



(a)



(b)

Figure 3.3: Comparison of the reference seismogram and initial model seismogram for the deep and shallow interval. Reference seismogram is the seismic signal obtained from the original well log and the initial model seismogram is obtained from the initial upscaled velocity seed. A Ricker source wavelet with a central frequency of 30 Hz is used to obtain both the seismograms

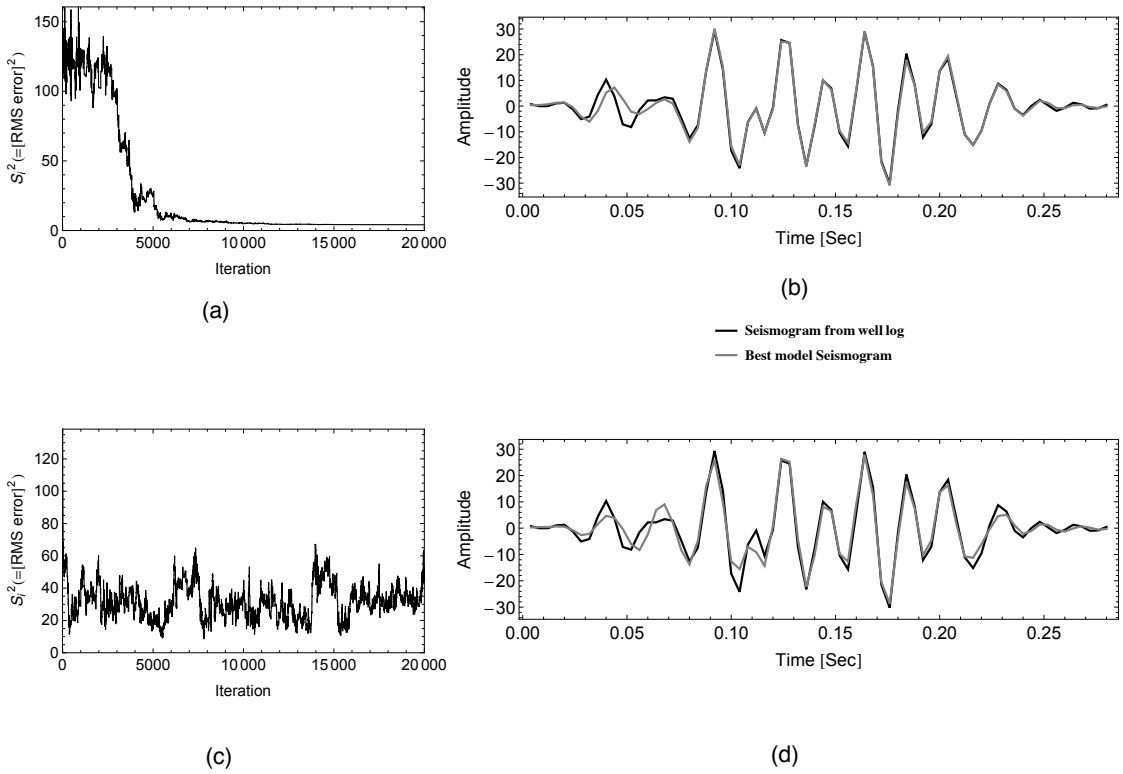


Figure 3.4: Plots in first row are obtained for Bayesian inversion while using tempered likelihood function for shallow interval. (a) Plot of mismatch between reference seismogram and model seismograms versus iterations (b) Comparing reference seismogram with best model seismogram. Plots in second row are obtained for Bayesian inversion without tempered likelihood function for shallow interval. (c) Plot of mismatch between reference seismogram and model seismograms versus iterations (d) Comparing reference seismogram with best model seismogram.

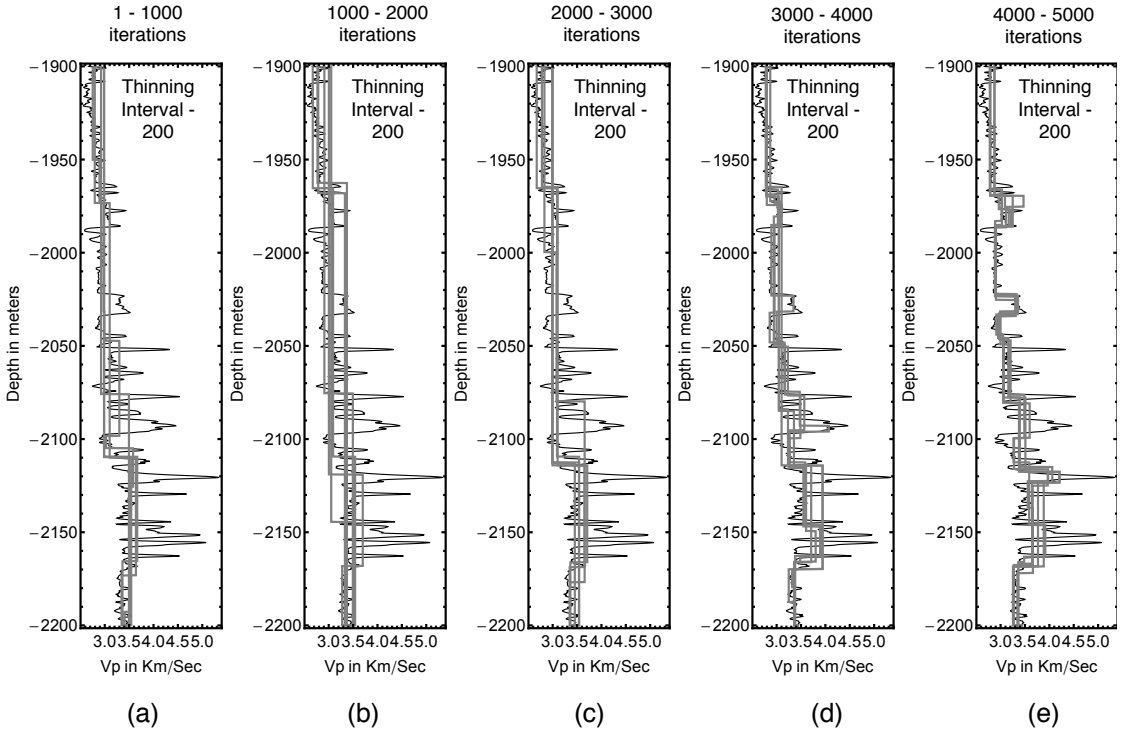


Figure 3.5: Ensemble of upscaled models for shallow interval from (a) 1-1000 (b) 1000-2000 (c) 2000-3000 (d) 3000-4000 and (e) 4000-5000 iterations. The thinning interval of upscaled models displayed for each of the intervals is 200 implying there are 6 upscaled models displayed for each interval at an interval of 200 iterations.

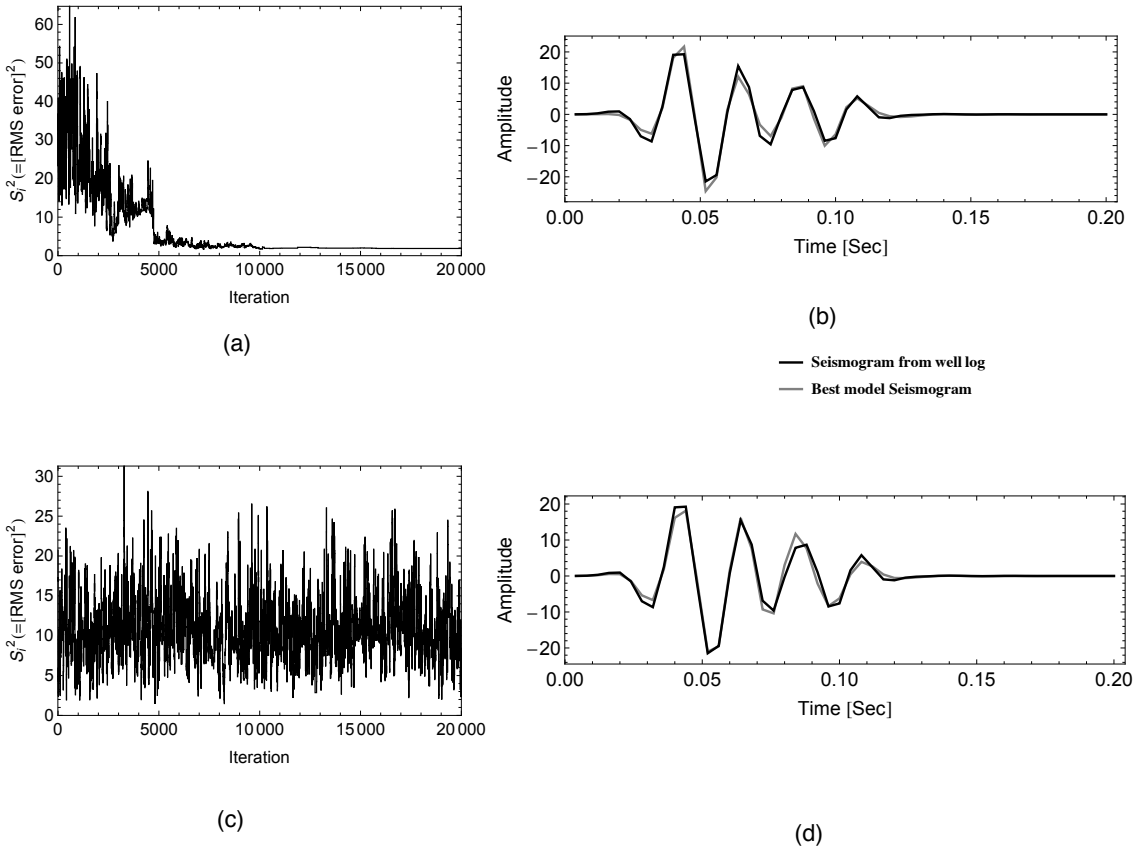


Figure 3.6: Plots in first row are obtained for Bayesian inversion while using tempered likelihood function for deep interval. (a) Plot of mismatch between reference seismogram and model seismograms versus iterations (b) Comparing reference seismogram with best model seismogram. Plots in second row are obtained for Bayesian inversion without tempered likelihood function for deep interval. (c) Plot of mismatch between reference seismogram and model seismograms versus iterations (d) Comparing reference seismogram with best model seismogram.

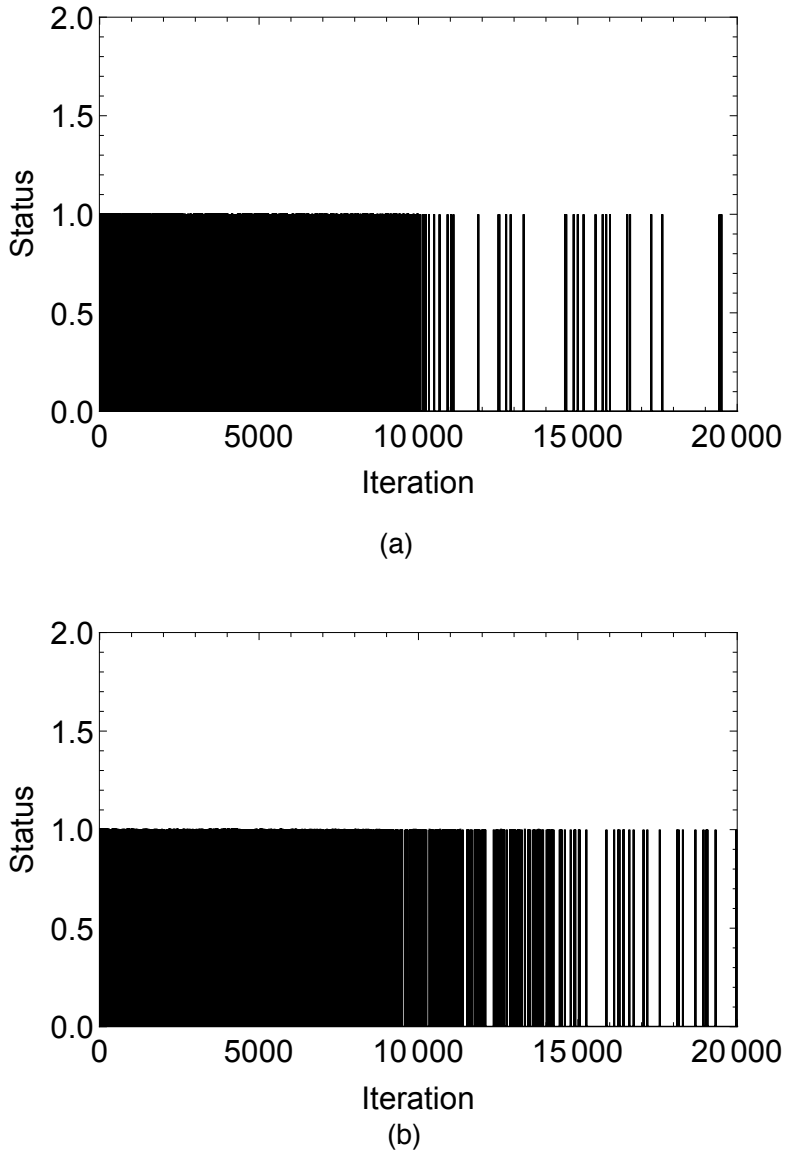


Figure 3.7: Acceptance status when running log blocking with rjMCMC Simulated Annealing for (a) deep and (b) shallow interval. All the accepted samples have an accepted status of 1 and those rejected have a value of zero.

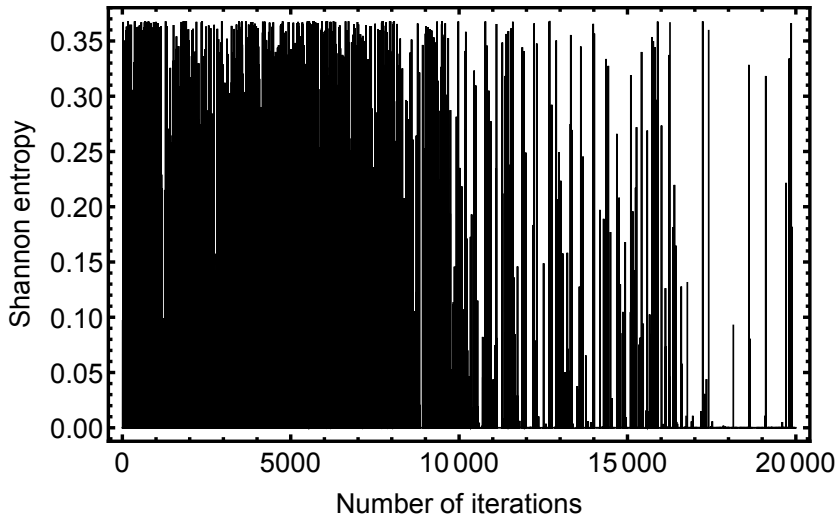


Figure 3.8: Shannon entropy measured from probability of upscaled models for the shallow interval during log blocking.

more of the layer boundaries are placed near a sharp contact in the velocity log will result in a reduction in the RMS error. Figures 3.4 and 3.6 compare the reference seismogram with the best model seismograms for both the intervals after log blocking with and without SA. Figures 3.7 shows the acceptance status for both the intervals at each iteration for the log blocking case with SA. Samples are accepted regularly in the first 5000 iterations shown as black color in the plots and subsequently most of the samples are rejected implied by the gaps in the accepted status plots.

In information theory, the uncertainty in acceptance of samples is quantified by Shannon's entropy. Shannon entropy is a measure of uncertainty in a random variable. The probability of the upscaled models during the random sweep can be used to calculate Shannon entropy. Entropy is given by

$$H(P) = -\sum_{i=1}^{n} P(\mathbf{m}_i) \log P(\mathbf{m}_i), \tag{3.22}$$

where $\mathbf{m}_i$ is the upscaled model at an iteration and n is the possible number of models. $P(\mathbf{m}_i)$ is the probability of an upscaled model and 1 - $P(\mathbf{m}_i)$ is the probability of all other possible models. Figure 3.8 shows entropy versus iterations during the Markov chain. The higher values of entropy during the initial iterations suggest all the models are equally probable compared to lower values of entropy after 10000 iterations implying some optimized models are identified.

To check if the rjMCMC sampler without SA is sampling from a stationary distribution, I compared the posterior distribution of number of layers in four Markov chains ran using well dispersed initial velocity models for the deep interval. Figure 3.9 shows similar histograms of number of layers from four Markov chains suggesting the results are sampled from a stationary distribution. The convergence assessment scheme explained earlier is used to determine the number of burn-in samples. The convergence assessment results discussed here are based on pooled Markov chain runs for log blocking of the deep interval. I show the results for a univariate case where samples from a single depth are interpreted to make inference about convergence. Figure 3.10 shows the PSRF's and individual plots of convergence parameters plotted together. Both PSRF's are close to 1 and below the cut off value of 1.2 mentioned by Brooks and Gelman (1998). $\hat{V}(\theta)$ and $W_c(\theta)$ converge and stabilize at a value of 0.0068 which suggests the variance of the posterior distribution of velocity at 2660 m is approximately 0.0068. Similarly $W_m(\theta)$ and $W_m W_c(\theta)$ values converge to 0.0066 which is an indication of variance of samples within models with same number of layers and is slightly lower than the variance from all the models. In the case of multivariate convergence analysis, Castelloe and Zimmerman (2002) proved $MPSRF$ i.e. the maximum eigenvalue of $W_c(\boldsymbol{\theta})^{-1}\hat{V}(\boldsymbol{\theta})$ is always greater than the maximum of $PSRF(\theta_i)$ (ratio of $\hat{V}(\theta)$ and $W_c(\theta)$ at individual depths). Figures 3.11 and 3.22 show the plots for convergence parameters from 5 different depths at the deep interval

after running log blocking in 4 Markov chains. The Maximum eigenvalues for the deep interval (Figures 3.11) are close to 1 but slightly higher than the $PSRF$ values at 2660 m shown in Figure 3.10. Results from the multivariate case suggest strong convergence early in the random sweep (less than 1000 sweeps out of 20000) and hence I used all the samples here for inference. A similar early convergence pattern is identified when running log blocking for the shallow interval and hence the results are not shown here.

Figures 3.13 and 3.12 show the histogram of layer boundaries from all the models for shallow and deep interval respectively with results from both the log blocking methods. Also shown in the figures are mean, lower and upper quantiles values of upscaled velocities from all the models. The standard deviation for the boundary depths, velocity and number of layers are higher in the case of log blocking without SA compared with log blocking with SA. Histograms of layer boundaries at all the contacts identified from the velocity log are narrower when using log blocking without SA compared to the wider distributions obtained after log blocking with SA. Similarly the different between lower and higher quantiles of upscaled velocity values, which is a measure of velocity uncertainty, is higher from samples obtained using log blocking without SA compared to corresponding values from log blocking with SA. This observation about uncertainty in layer boundary depths and velocity is true for both the intervals at all the depths. In Figure 3.14, I compare the standard deviation of impedance values in all the iterations at 2640, 2680 and 2720 m of the deep interval respectively. Both the approaches traverse the model space (the prior range of impedance at the corresponding depth) as observed from Figure 3.14, however, the wider distributions of histograms obtained after using log blocking without SA generate the true estimate of uncertainty for the upscaled velocities. To demonstrate how the velocities change with iterations during the random sweep, I

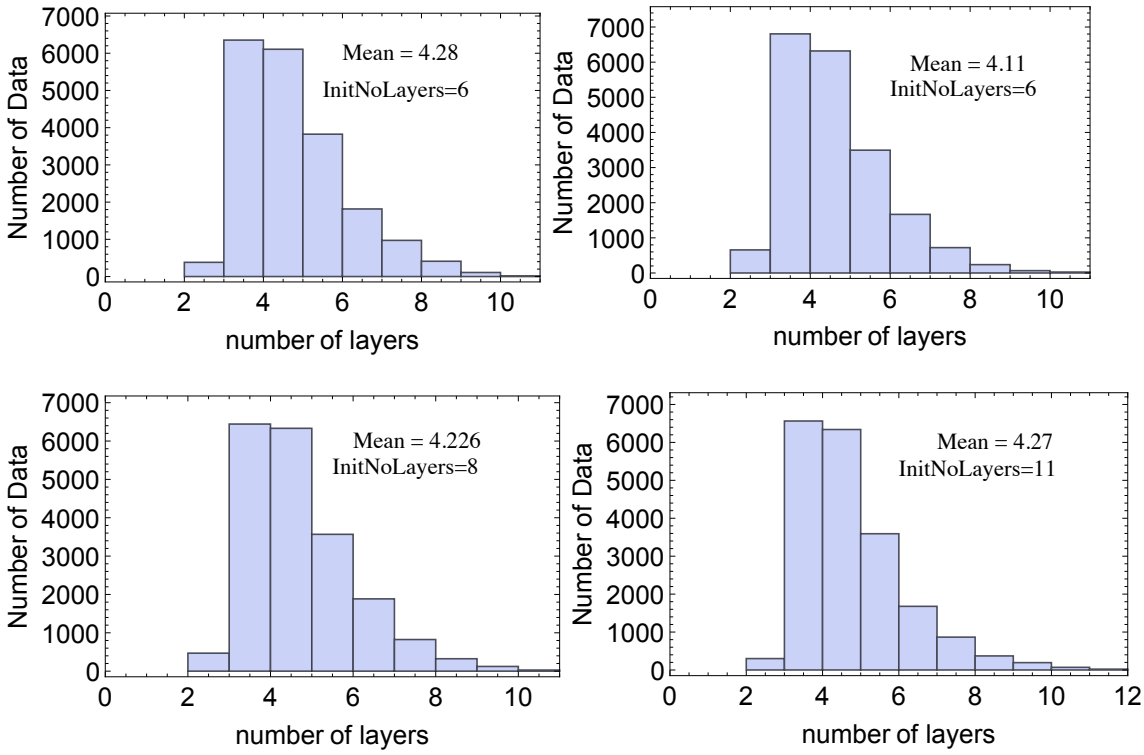


Figure 3.9: The histogram of number of layers from all the four chains while running log blocking on velocity log from deep interval using zero-offset seismogram as reference seismogram. Also shown are the mean number of layers and the initial seed number of layers.

showed the plots of the Markov chain for velocity at 2720 m and layer boundary nearest to 2720 m in the deep interval in Figure 3.15.

3.4 Log blocking using multi-offset seismograms

In this section, I extend log blocking to obtain upscaled velocity models when using multi-offset seismograms as reference seismogram. These velocity models can be used as rough initial guess models at the well location for full waveform inversion while using pre-stack seismic data as reference data. The optimal upscaled velocity models can also be used to generate petrophysical properties at the scale of grid cells used in reservoir flow simulations by applying appropriate rock physics relationships.

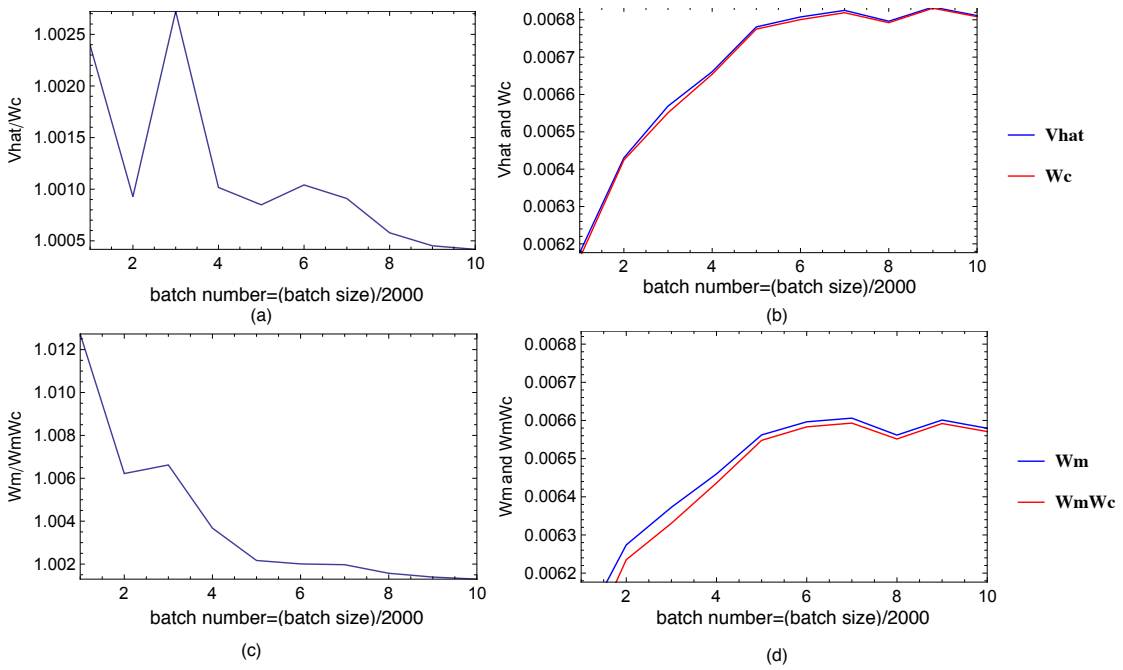


Figure 3.10: Convergence parameters calculated from the Markov chain of upscaled velocity at 2660 m of the deep interval. The results here are for univariate convergence parameters since they are derived for a single model parameter i.e. velocity at a certain depth. $\hat{V}(\theta)$ can be interpreted as variance of samples in all the chains, $W_c(\theta)$ is the average of sample variance within each chain implying it is a measure of within chain variance, $W_m(\theta)$ is the variance of samples within models with same number of layers from all the chains. $W_m(\theta)W_c(\theta)$ is the variance of models with same number of layers within each chain.

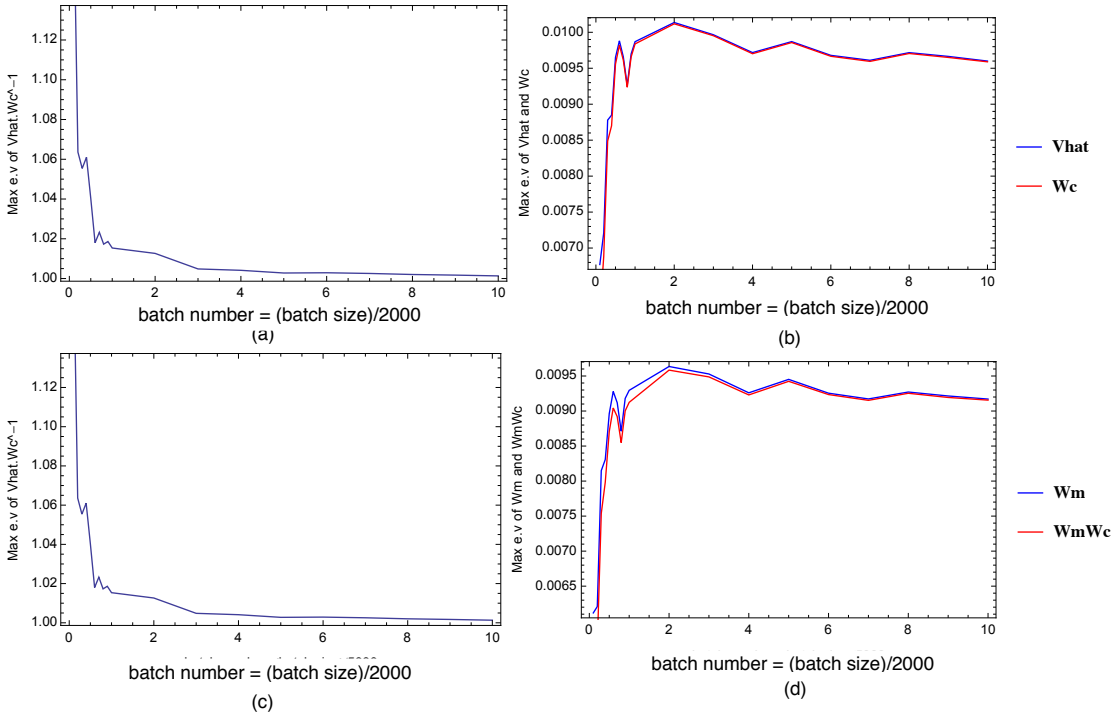


Figure 3.11: Multivariate convergence parameters calculated from the Markov chains of upscaled velocities at 2620, 2640, 2660, 2680 and 2700 m of the deep interval. $\hat{V}(\theta)$ can be interpreted as variance of samples in all the chains, $W_c(\theta)$ is the average of sample variance within each chain implying it is a measure of within chain variance, $W_m(\theta)$ is the variance of samples within models with same number of layers from all the chains. $W_m(\theta)W_c(\theta)$ is the variance of models with same number of layers within each chain.

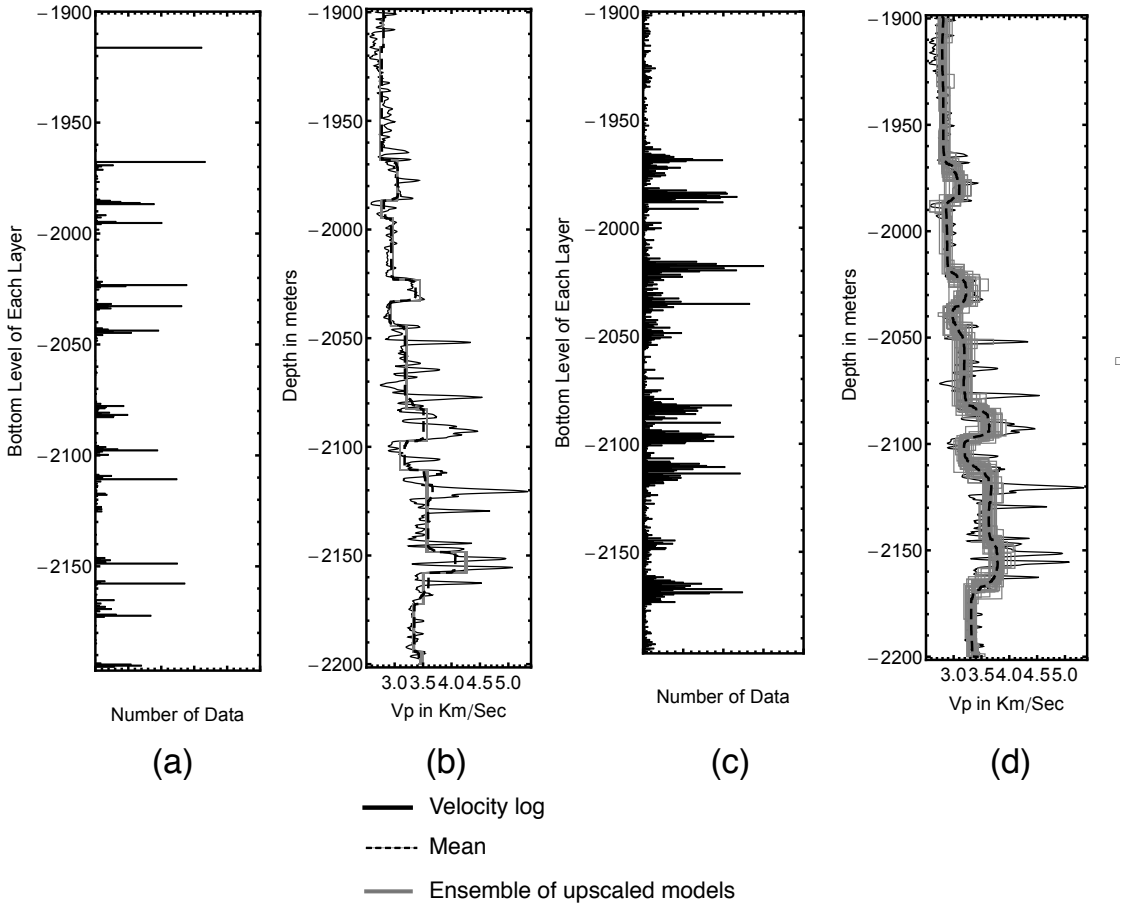


Figure 3.12: (a) Histogram of layer boundary depths for shallow interval and (b) Ensemble of 50 upscaled models extracted at a thinning interval of 100 models from 5000 best models generated using rjMCMC Simulated Annealing (c) Histogram of layer boundary depths and (d) Ensemble of 50 upscaled models extracted at a thinning interval of 100 models from 5000 best models generated using rjMCMC without Simulated Annealing.

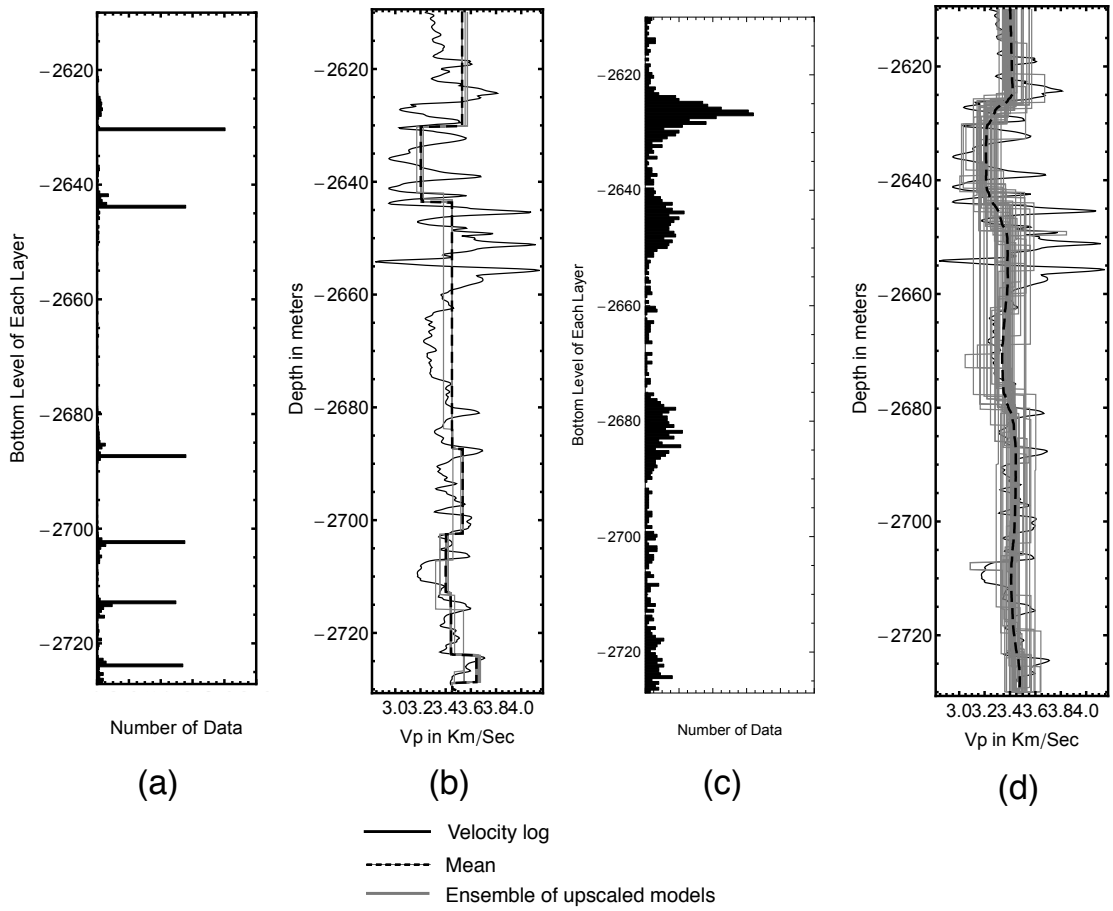


Figure 3.13: (a) Histogram of layer boundary depths for deep interval and (b) Ensemble of 50 upscaled models extracted at a thinning interval of 100 models from 5000 best models generated using rjMCMC Simulated Annealing (c) Histogram of layer boundary depths and (d) Ensemble of 50 upscaled models extracted at a thinning interval of 100 models from 5000 best models generated using rjMCMC without Simulated Annealing.

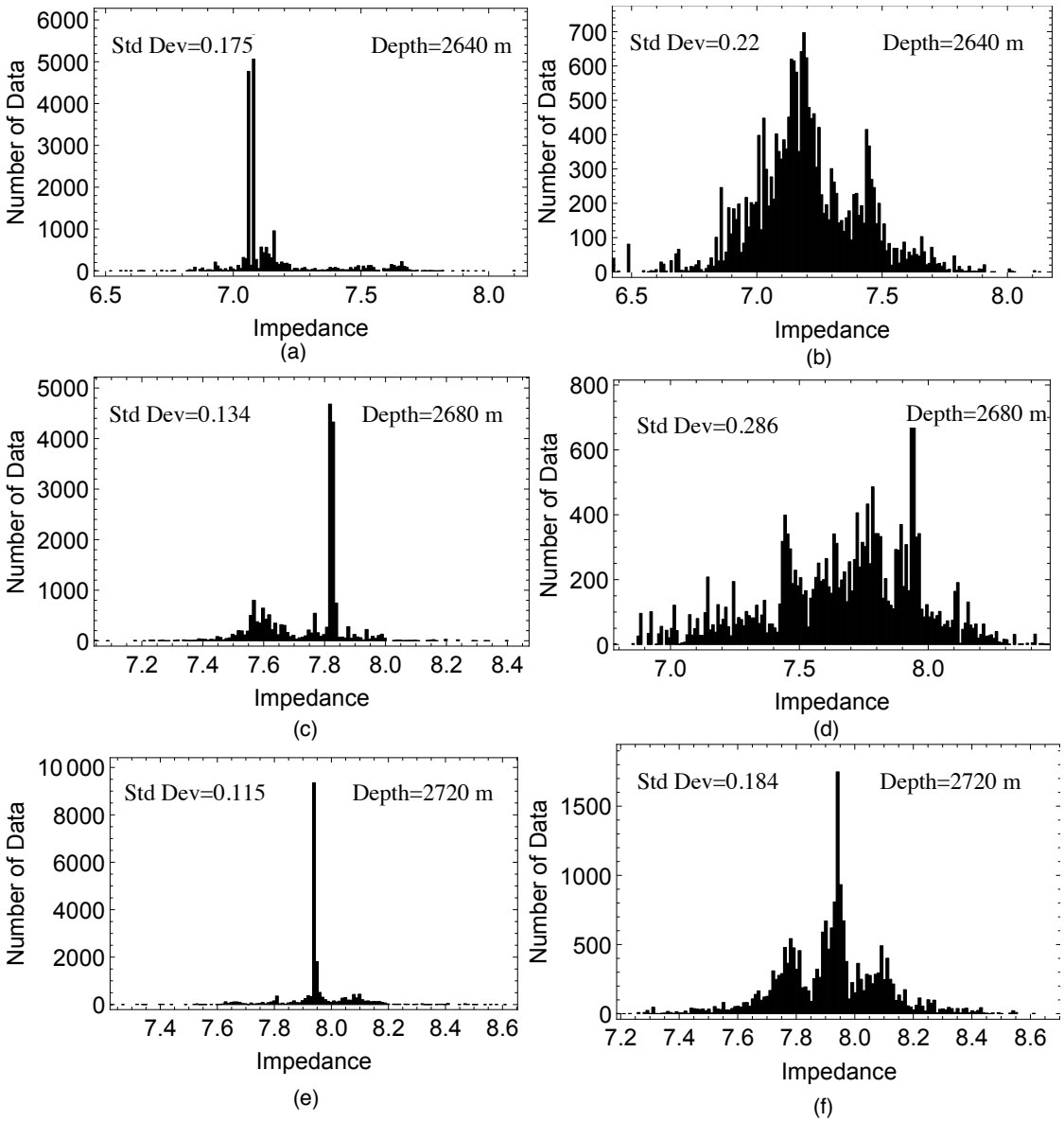


Figure 3.14: Comparison of impedance values while upscaling well logs when using rjMCMCSA and without SA. The first column shows impedance histograms when using rjMCMCSA at depths 2640, 2680 and 2720 m respectively. The second column shows impedance histograms when using rjMCMC without SA at depths 2640, 2680 and 2720 m respectively.

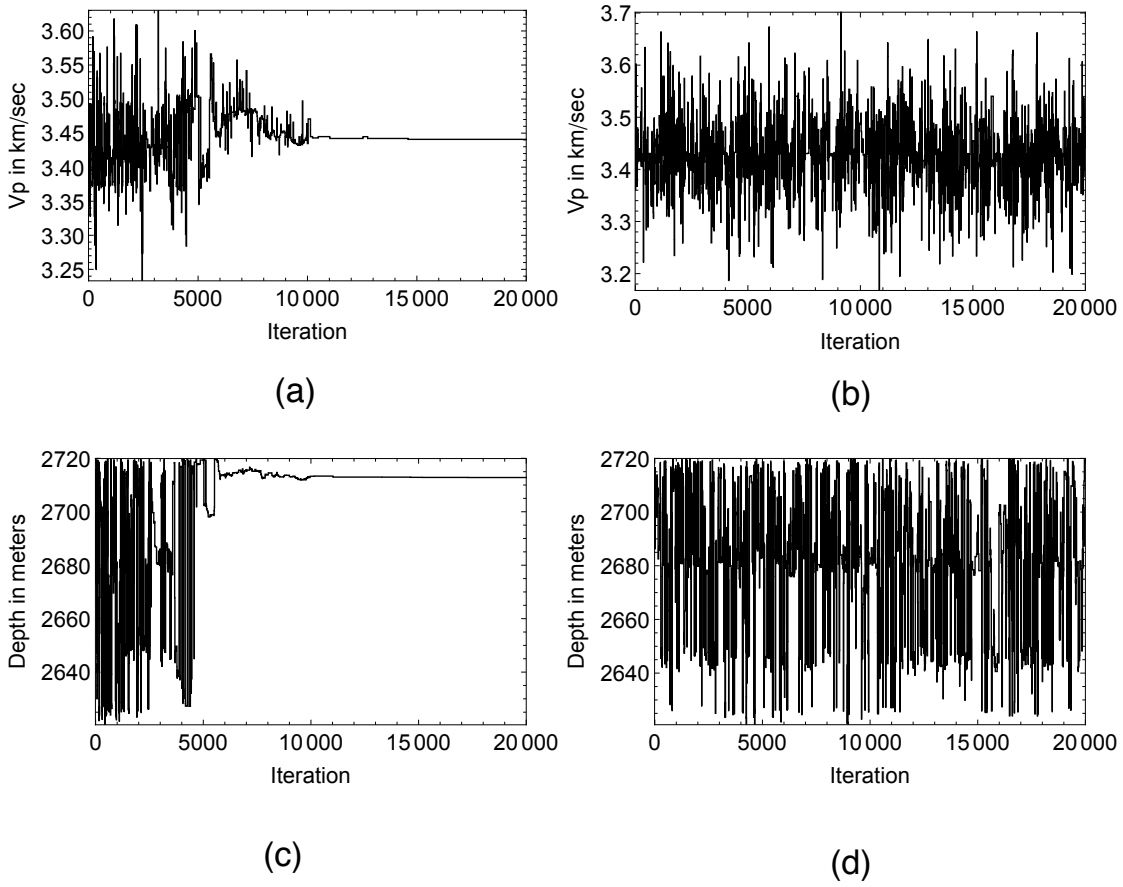


Figure 3.15: Velocity and nearest boundary depth versus iterations at 2720 m in deep interval. The first column shows the results when using rjMCMCSA and second column shows results when using rjMCMC without SA.

The objective of this section is to compare inversion results (upscaled velocity models) obtained when using zero-offset and multi-offset seismograms generated from velocity log as reference seismograms.

I use a 170 m log interval drilled in shallow sediments and basalts at location U1347 on TAMU Massif for log blocking using multi-offset seismograms as reference seismogram. TAMU massif is a large volcano at Shatsky Rise oceanic plateau. The velocity log at U1347 and the upper and lower limits of the prior distribution are shown in Figure 3.16. The initial upscaled velocity model is shown in Figure 3.16 and Markov chain design parameters (proposal distribution standard deviation, coefficient of walk size and standard deviation of error function) are the same while running log blocking using zero-offset and multi-offset seismograms as reference seismograms. The reference multi-offset seismogram consists of four seismograms generated in a fine layered media applying propagator matrix method using a 30 Hz source frequency Ricker wavelet. The four seismograms correspond to 0, 500, 1000 and 1500 m offsets. Figure 3.17 shows the reference multi-offset seismogram and the initial seismogram generated from the initial upscaled velocity model. The results generated in this section are obtained when using log blocking without any tempered likelihood function.

To verify if the rjMCMC sampler is sampling from a stationary distribution, I ran log blocking in four Markov chains and compared the frequency of number of layers visited. The number of layers in the four Markov chains are randomly chosen between 5 and 30 and the layer boundaries are also randomly chosen. Figures 3.18 and 3.19 show the RMS misfit^2 versus iterations plots and number of layer histograms from the four Markov chains respectively when using zero-offset seismograms as reference seismogram. The number of layers for the initial models are mentioned in Figure 3.19. Figures 3.20 and 3.21 show the RMS misfit^2 versus iterations plots and number of layer histograms from the four Markov chains respectively when

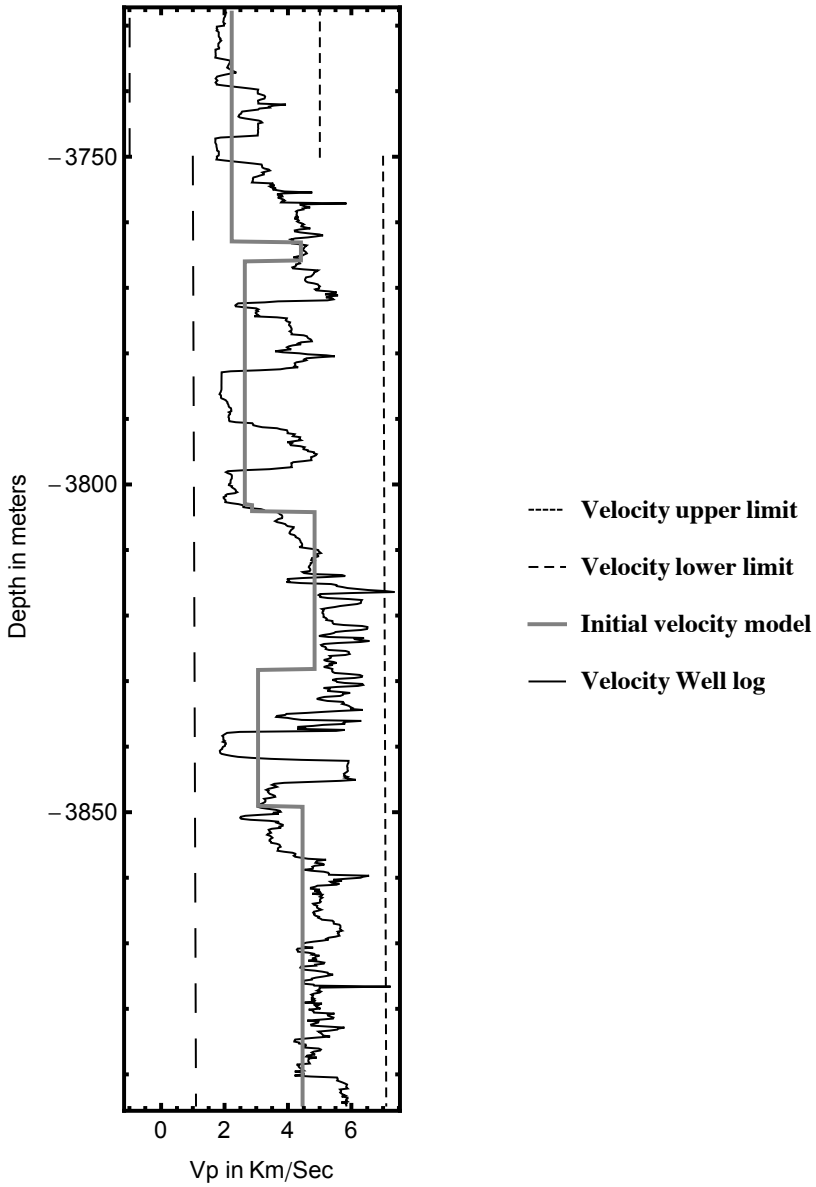


Figure 3.16: Initial upscaled velocity profile obtained from upscaling log measurements at U1347 location on Shatsky Rise oceanic plateau. The initial model is the same for upscaling velocity logs using zero-offset and multi-offset seismogram as reference seismograms.

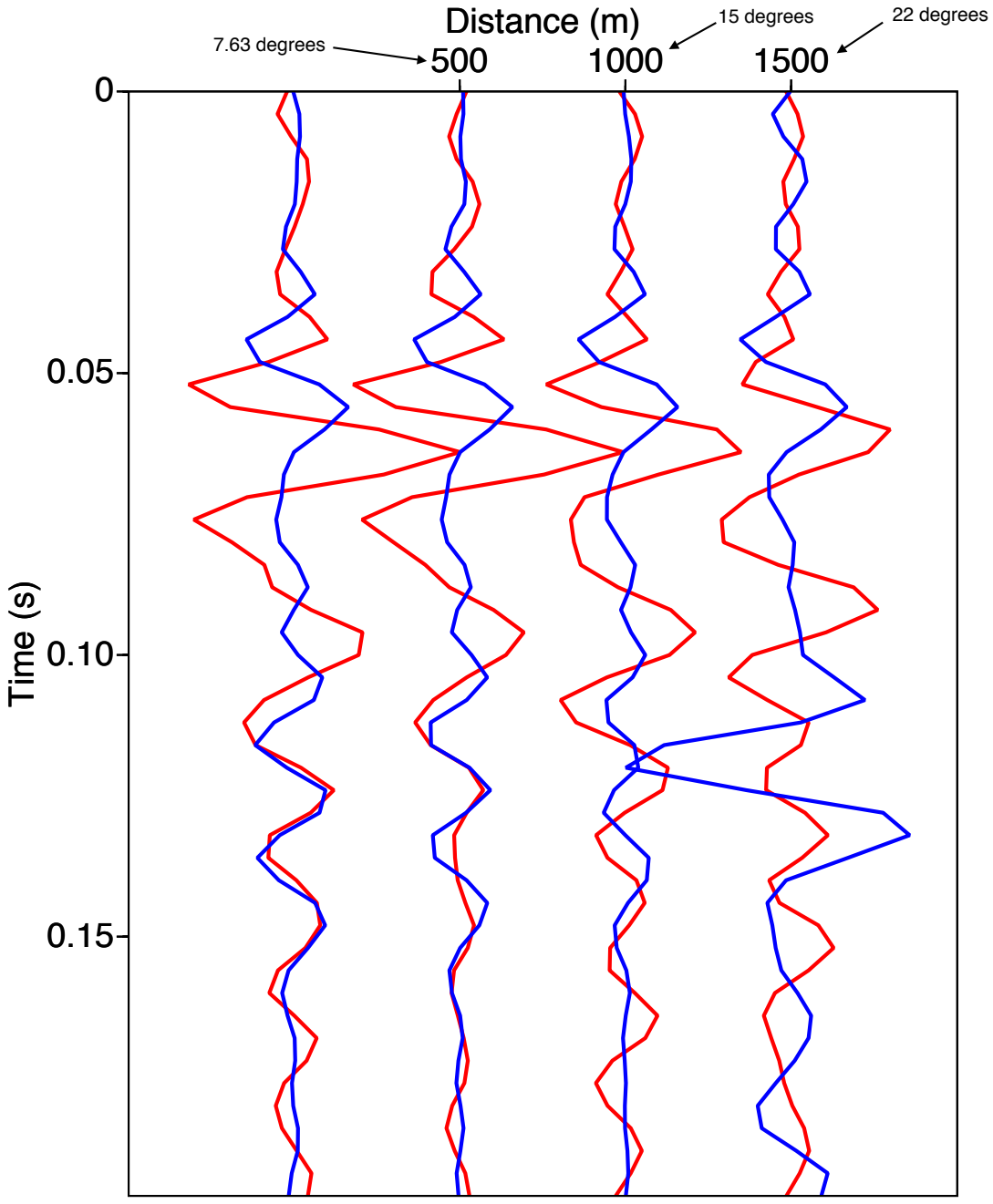


Figure 3.17: Comparison of the reference seismogram (red) at U1347 with the seismogram (blue) obtained from the initial upscaled velocity model.

using multi-offset seismograms as reference seismogram. In both the cases when zero-offset and multi-offset seismograms are used as reference seismograms, the inter chain histogram distributions of the number of layers are similar. The mean value of number of layers when using zero-offset and multi-offset reference seismograms are around 10 and 12 respectively. This suggests the chains are most likely sampling the models from the same stationary distribution. An appropriate design of Markov chain is essential to verify convergence before drawing inference. For example, the standard deviation of the proposal distribution is a design parameter for setting up Markov chain which effects the rate at which the Markov chain reaches convergence.

The convergence results discussed here are from log blocking of U1347 log interval using multi-offset seismograms as reference seismogram. $MPSRF$ values from U1347 log interval are only slightly less than 1.2 (Figure 3.22) compared to those at deep interval which are much closer to 1 implying a strong converge at deep interval and moderate converge at U1347. Only models sampled from 10000 to 50000 iterations should be used for inference in the case of log blocking for U1347 log interval since the $MPSRF$ is less than 1.2 at approximately 10000 iterations (batch number 2 in Figure 3.22). $\hat{V}(\boldsymbol{\theta})$ and $W_c(\boldsymbol{\theta})$ calculated in multivariate convergence assessment does not have any physical meaning like in the univariate case where $\hat{V}(\theta)$ and $W_c(\theta)$ are measurements of total variation of samples. However, $\hat{V}(\boldsymbol{\theta})$ and $W_c(\boldsymbol{\theta})$ can be used to compare uncertainty in the posterior distribution of upscaled velocities for two different logs chosen for convergence assessment. For example, the maximum eigenvalue of $\hat{V}(\boldsymbol{\theta})$ and $W_c(\boldsymbol{\theta})$ at U1347 on Shatsky Rise converge to 0.23 (Figure 3.22) which is higher than the corresponding value of 0.0095 for the deep interval. The number of burn-in samples while running log blocking with zero-offset seismogram as reference seismogram are also close to 10000 samples. The convergence parameter results for this case are not shown here as there are similar to the case when multi-offset

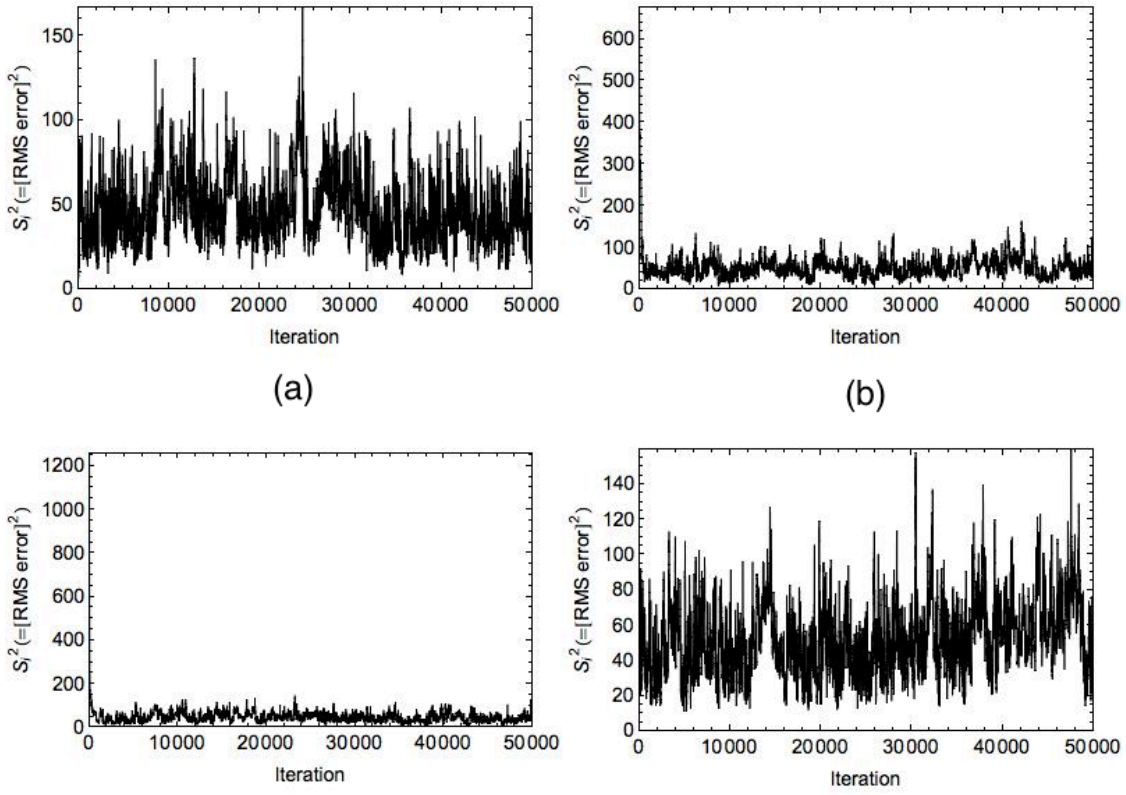


Figure 3.18: Mismatch between the reference and model seismograms while running log blocking on velocity log from U1347 with four different initial seeds. The reference seismogram is the zero-offset seismogram. The location of the boundary depths and number of layers are different in the four initial seeds.

seismogram is used as reference seismogram.

All the sharp contacts between volcanoclastics and basalts layers are identified with narrow distribution in the layer boundary histogram. Figures 3.23 and 3.24 show the histogram of layer boundaries and the mean, higher and lower quantiles of upscaled velocity for the log interval at U1347 while using normal incidence and multi-offset seismograms as reference seismograms respectively. The layer histogram plots show ∼10 boundary distributions are required to represent the upscaled medium

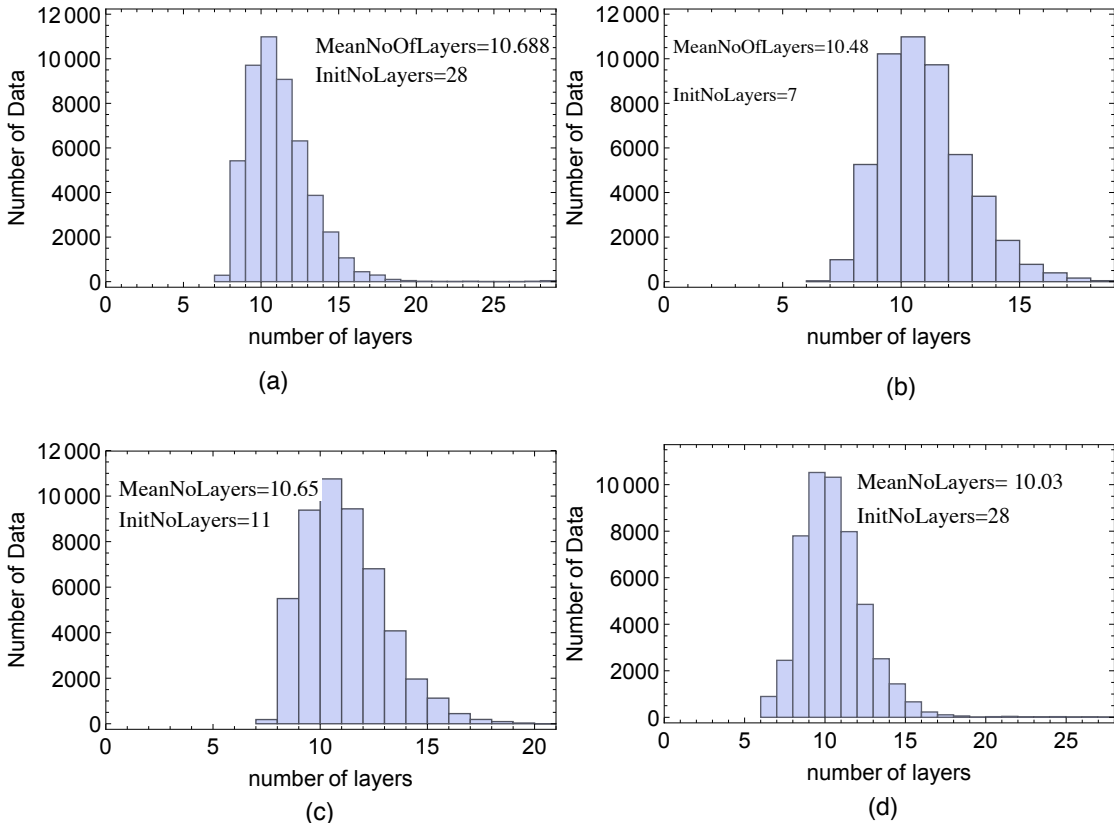


Figure 3.19: The histogram of number of layers from all the four chains while running log blocking on velocity log at U1347 using zero-offset seismogram as reference seismogram. Also shown are the mean of the number of layers posterior distribution and initial number of layers.

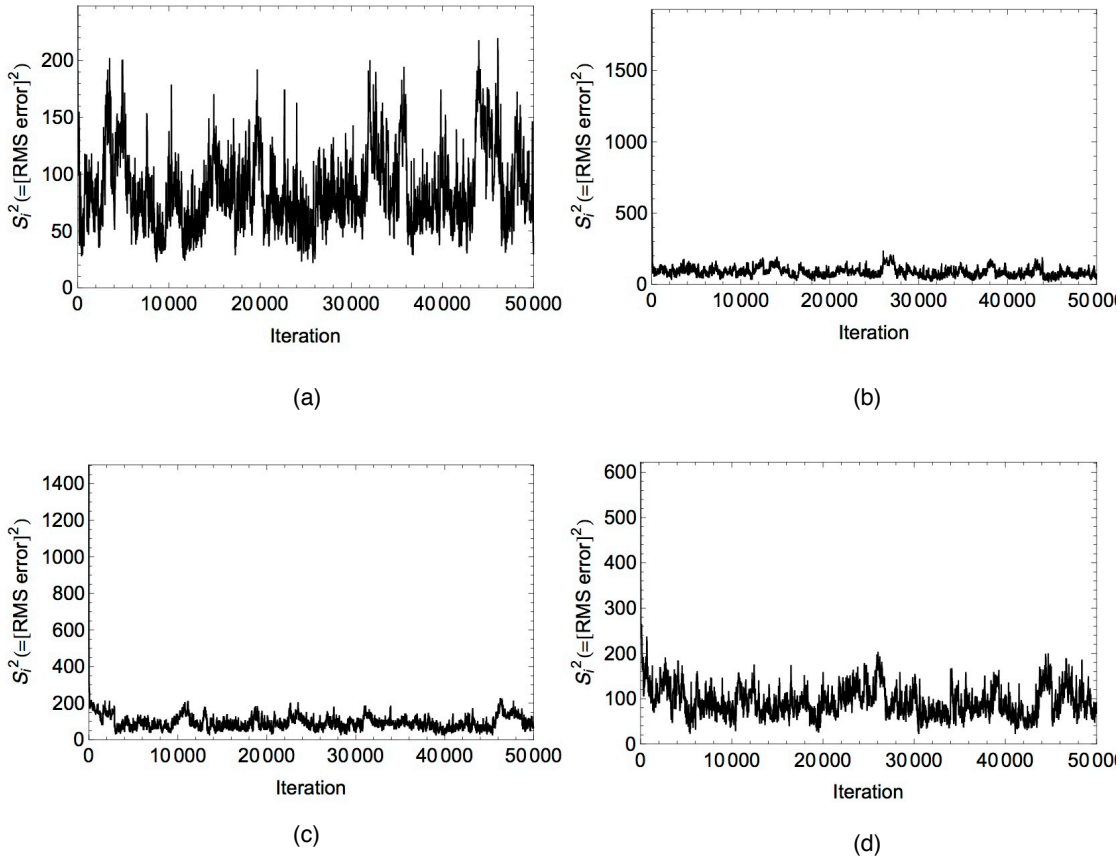


Figure 3.20: Mismatch between the reference and model seismograms while running log blocking on velocity log from U1347 with four different initial seeds. The reference seismogram is the multi-offset seismogram. The location of the boundary depths and number of layers are different in the four initial seeds.

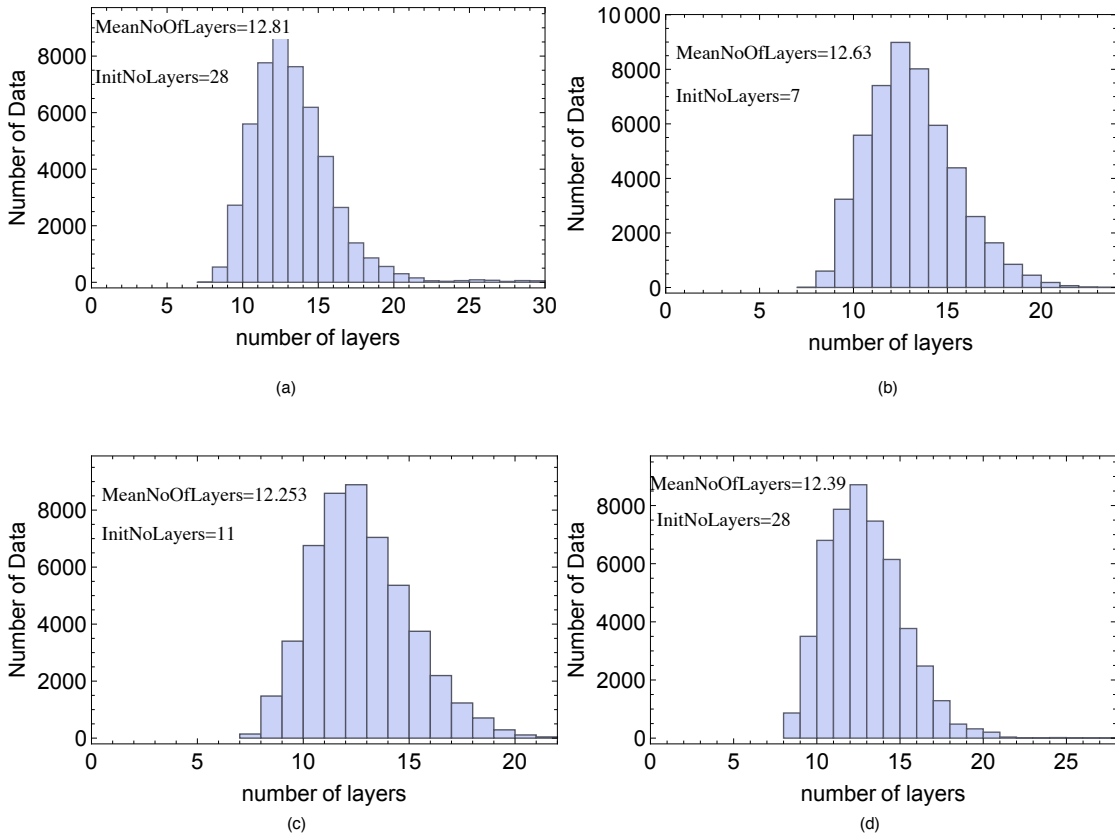


Figure 3.21: The histogram of number of layers from all the four chains while running log blocking on velocity log at U1347 using multi-offset seismograms as reference seismogram. Also shown are the mean of the number of layers posterior distribution and initial number of layers.

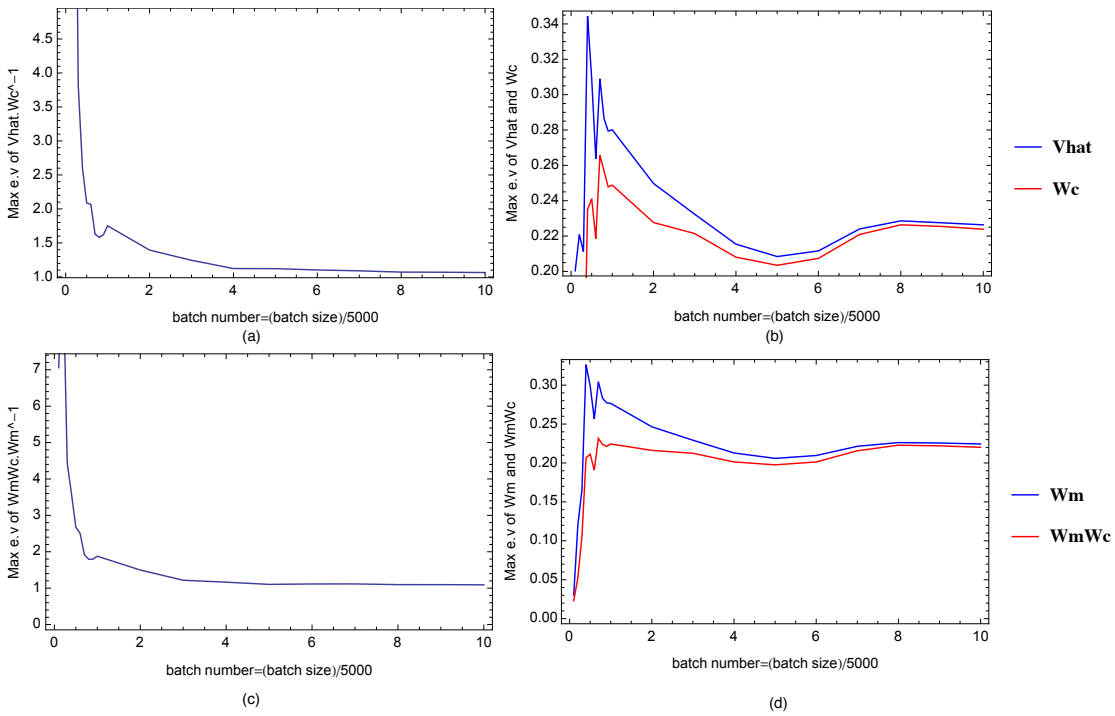


Figure 3.22: Multivariate convergence parameters calculated from the Markov chains of upscaled velocities at 3750, 3775, 3800, 3825 and 3850 m of the U1347 velocity log. $\hat{V}(\theta)$ can be interpreted as variance of samples in all the chains, $W_c(\theta)$ is the average of sample variance within each chain implying it is a measure of within chain variance, $W_m(\theta)$ is the variance of samples within models with same number of layers from all the chains. $W_m(\theta)W_c(\theta)$ is the variance of models with same number of layers within each chain.

compared to ∼12 when using multi-offset seismograms as reference seismogram. For example, the layer distributions at ∼3800 m and ∼3740 m in Figure 3.24 are missing in Figure 3.23. Figure 3.25 compares the number of layer histograms and the mean value of number of layers when using multi-offset seismogram as reference seismogram to the case when using normal incidence as reference seismogram. Presence of more layer distributions when running log blocking with multi-offset seismograms as reference seismogram suggest the need for a more complex upscaled medium compared to the case when normal incidence is the reference seismogram. The mean velocity of all the upscaled models when using normal incidence as reference seismogram is smoother and the uncertainty in upscaled velocity (difference between upper and lower quantiles) is higher compared to the upscaled results when using multi-offset seismogram as reference seismogram as shown in Figures 3.23 and 3.24. Figure 3.26 compares the reference seismograms with the best model seismograms.

3.5 Conclusions

Log blocking when run with a tempered likelihood function assists in identifying the optimal models whereas log blocking without a tempered likelihood function estimates the uncertainty of the full posterior distribution of model parameters. The log blocking procedure mentioned here provides a workflow to generate upscaled elastic properties and quantify the uncertainty in upscaled elastic properties. The technique uses the standard Backus average and allows sufficient deviation of layer velocities from Backus average to generate model seismograms which are similar to the full waveform generated from the fine layer media. The convergence assessment for rjMCMC sampling identifies a point in the Markov chain when the chain has converged and assists in drawing inference of the full posterior distribution.

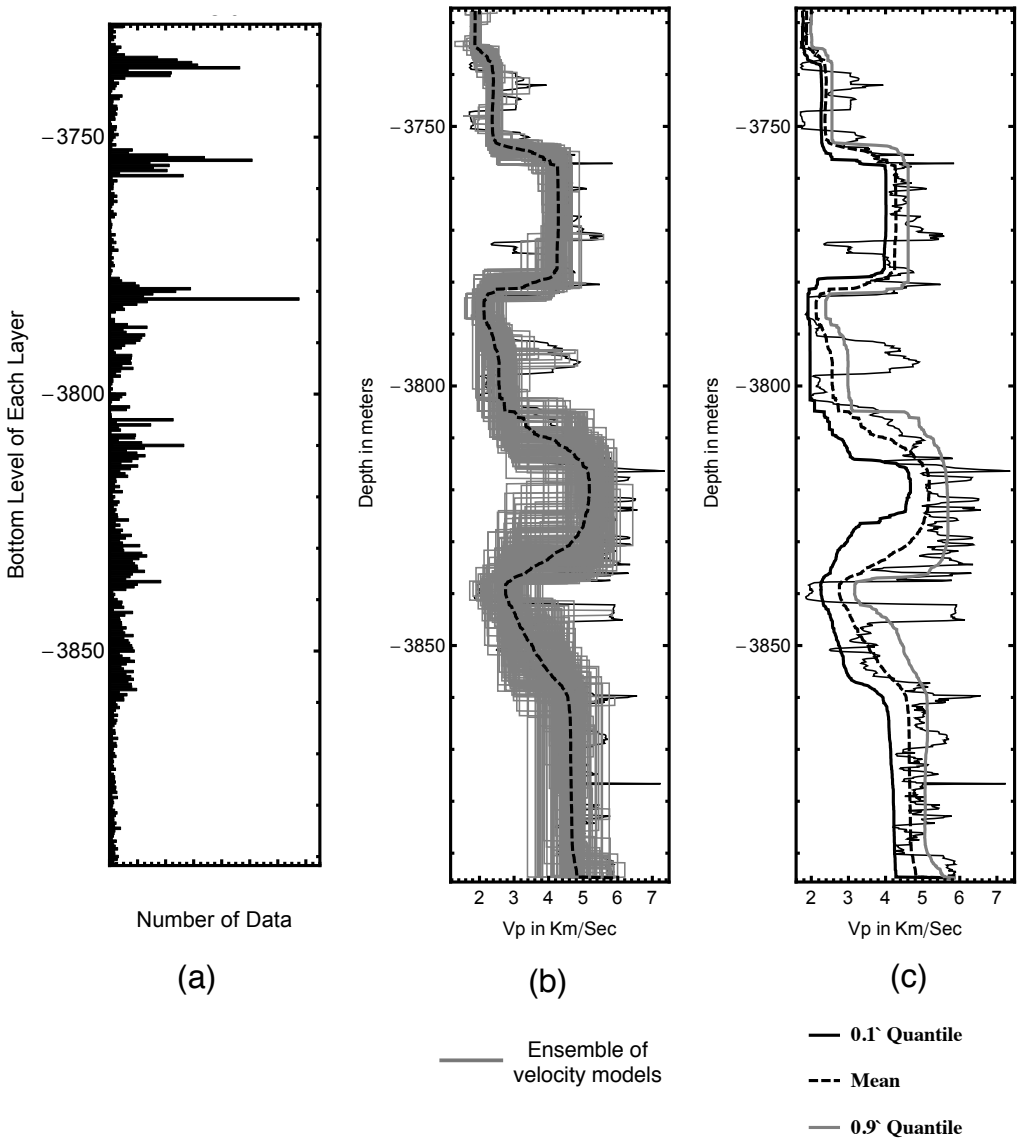


Figure 3.23: (a) Histogram of layer boundary depths and (b) Ensemble of 200 upscaled models extracted at a thinning interval of 200 iterations from 40000 iterations for log interval measured at U1347 while running rjMCMC sampling for 50000 iterations using zero-offset seismogram as reference seismogram. (c) Mean, upper and lower quantiles of 40000 upscaled velocity models after removing 10000 burn-in samples.

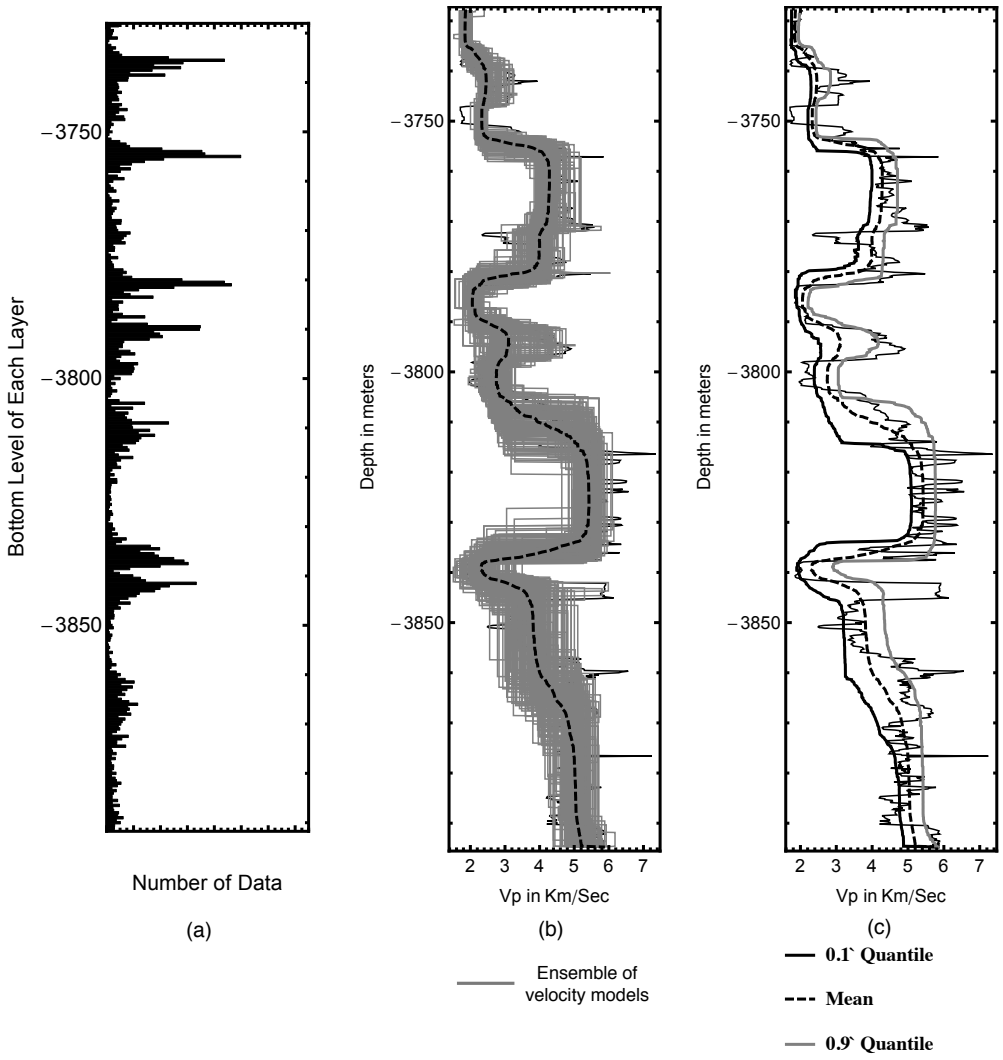


Figure 3.24: (a) Histogram of layer boundary depths and (b) Ensemble of 200 upscaled models extracted at a thinning interval of 200 iterations from 40000 iterations for log interval measured at U1347 while running rjMCMC sampling for 50000 iterations using multi-offset seismogram as reference seismogram. (c) Mean, upper and lower quantiles of 40000 upscaled velocity models after removing 10000 burn-in samples.

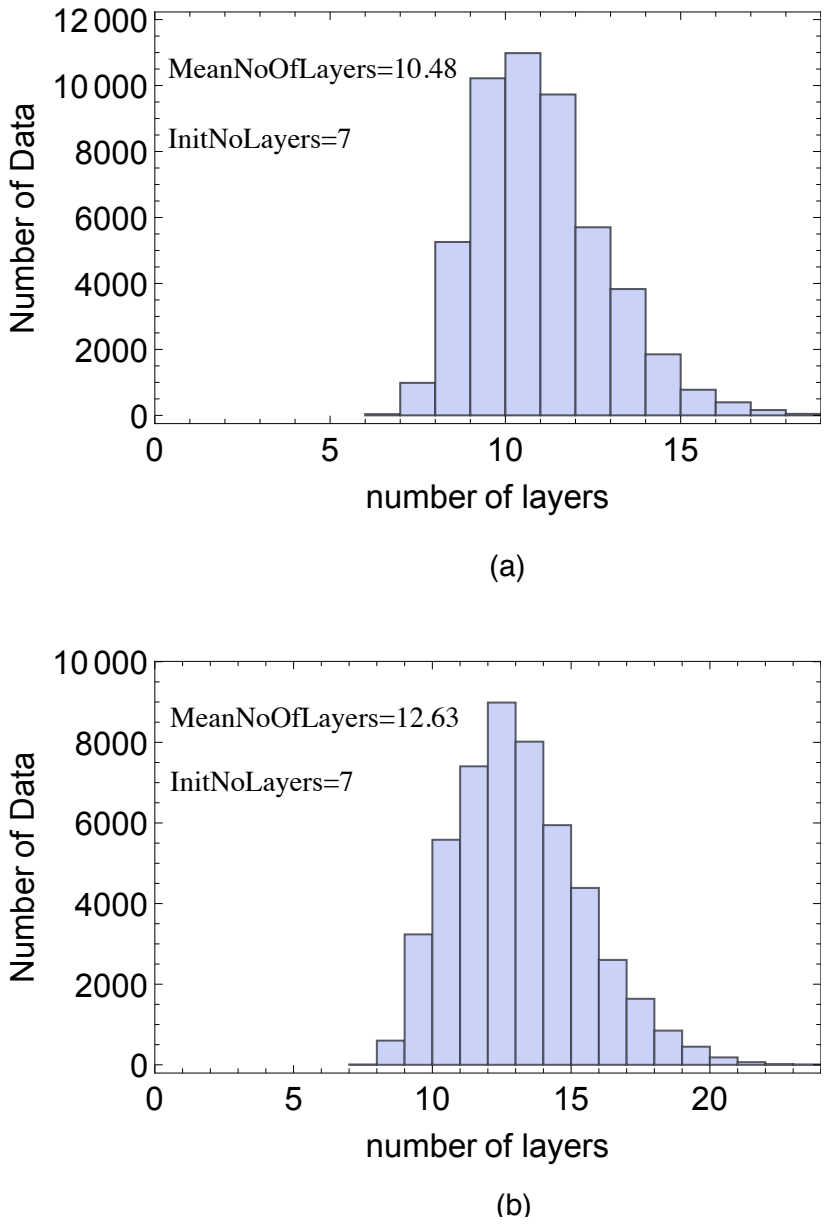
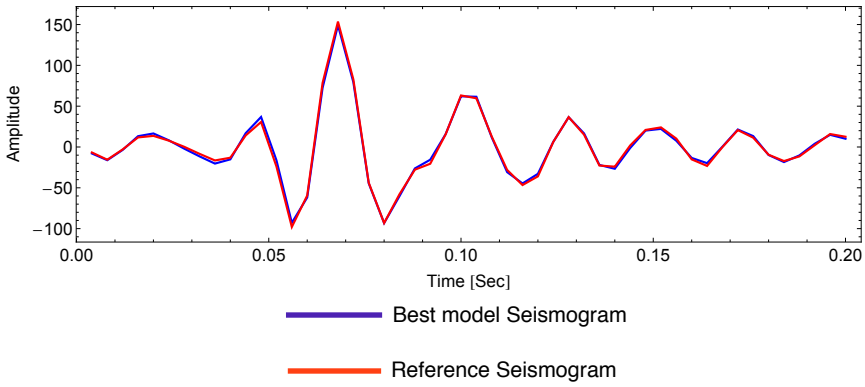
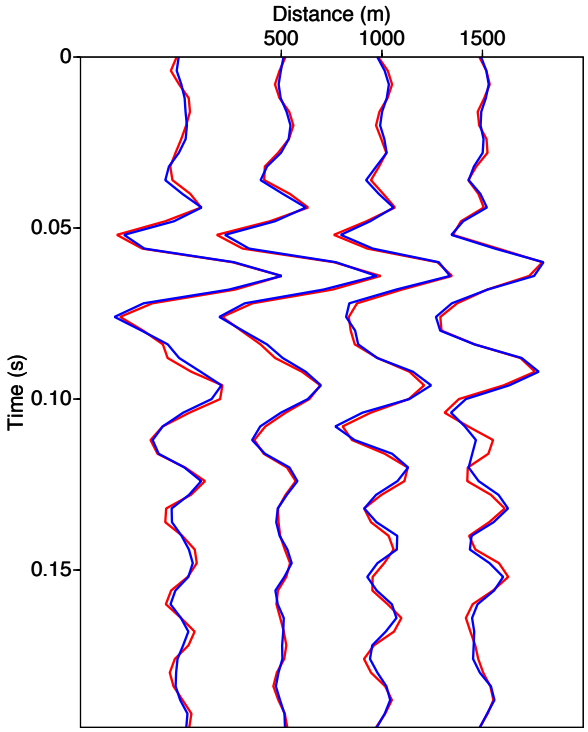


Figure 3.25: Number of layers histogram after running Markov chain for upscaling velocity log at U1347 by using (a) normal incidence seismogram as reference seismogram (b) multi-offset seismogram as reference seismogram.



(a)



(b)

Figure 3.26: Comparison of the reference seismogram (red) at U1347 with the seismogram (blue) obtained from the best upscaled velocity model when using (a) normal incidence seismogram as reference seismogram (b) multi-offset seismogram as reference seismogram.

4. ESTIMATION OF SHALLOW IMPEDANCE STRUCTURE AT SHATSKY RISE TO STUDY THE LATE-STAGE EVOLUTION OF TAMU AND ORI VOLCANOES

4.1 Introduction

Shatsky Rise is an oceanic plateau located in the northwest Pacific Ocean which is formed as a result of massive eruption of lava flows. Evolution of oceanic plateaus is poorly understood due to their remote submarine locations and identifying them on the seafloor is a challenging task due to the gentle slope of the large volcanoes comprising the oceanic plateaus. The volume of lava flows comprising the oceanic plateaus and continental flood basalt provinces (CFBP) suggest large sources for lava. There are multiple explanations for the source mechanisms of lava building these plateaus. The mantle plume hypothesis suggests the source of lava is from a deep mantle plume from lower mantle which causes uplift of the lithosphere and erupts huge volume of lava (Richards et al., 1989; Mahoney et al., 1993; Duncan and Richards, 1991; Coffin and Eldholm, 1994). This mechanism is widely accepted for source mechanism for oceanic plateaus. However at Shatsky Rise, the geochemistry of the rocks suggest they are of Mid-Ocean Ridge (MORB) origin whereas the geochemistry should resemble that of lower mantle if Shatsky Rise is formed from a mantle plume. Another mechanism is a possible decompression of upper mantle below the lithosphere and the drawback for this explanation is if the source volume would be large enough to feed the large oceanic plateaus (Foulger, 2007). The last possible mechanism is due to meteorite impacts but no evidence of any meteorite rocks raises questions about this hypothesis (Ingle and Coffin, 2004). Similar oceanic plateaus include Ontong Java Plateau, Caribbean-Colombian Plateau and Manihiki Plateau and CFBP's are

Columbia River Basalts and Parana Flood Basalts.

2D Multichannel seismic (MCS) reflection lines are acquired on Shatsky Rise and other Oceanic plateau to draw clues about the evolution of the oceanic plateaus. Scattering of energy due to internal reflections within basalts and the high impedance contrast between basalts and shallow sea sediments result in weak reflection below the first major basement reflection. The objective of this chapter is to perform impedance inversion from post-stack seismic data to study the shallow impedance structure at Shatsky Rise and draw inference about the late-stage evolution of the TAMU and Ori Volcanoes in the Shatsky Rise oceanic plateau. Impedance can be used to characterize lithology by assigning appropriate cut offs to identify lithologies. To invert for impedance from the weak reflections requires an inversion scheme which can use prior information and predict the impedance accurately. The intra basalt reflections are due to impedance contrast, between different basalt types and, between basalt and volcanoclastic sediments (Zhang et al., 2012). The thickness of the basalt and sediment layers vary from few meters to tens of meters and this variation in thickness poses a challenge to perform fixed dimensional stochastic inversion to obtain impedance values. For example, to perform impedance inversion of a 1 km interval would require using at least 50 model parameters assuming the thinnest layer thickness is 20 m. To avoid this problem, I use a transdimensional approach for seismic inversion where the number of layers, layer boundaries and impedance values are defined as model parameters. Using the reversible jump MCMC approach reduces the computational cost and avoids curse of dimensionality which is noticed in high-dimensional inversion problems. The vertical resolution of the impedance results is limited by the bandwidth of the post-stack seismic data. The seismic model is also dependent on the initial low frequency impedance model. The typical vertical resolution during seismic inversion is approximately about $\lambda/4$ (λ is

the wavelength). $\lambda/4$ value for a wave traveling in a medium with p-wave velocity of 4 km/sec (approximate value for upper igneous crust) generated with a source frequency of 30 Hz is ∼44 m.

In this chapter, I will perform transdimensional inversion to obtain 2D impedance profiles from the summit to the flank of the TAMU (52 km in length by ∼2 km depth) and Ori (80 km in length by ∼2 km depth) Massifs using the MCS reflection surveys. For calibration, I compare the inverted impedance values with the log measurements at an IODP drill site located on TAMU Massif. In addition, I compare the percentage of basalt types from drill cores with those inferred from inverted impedance values from seismic data. The three lithologies identified in the drill cores include Massive basalt flows, pillow basalts and volcanoclastics. The 2D impedance results from TAMU and Ori Massif suggest the amount of pillow or geochemically altered basalts is higher at Ori Massif compared to that of TAMU Massif. There are more pillow basalts at the summit compared to that at the flank of TAMU Massif.

Below I discuss the geologic history of Shatsky Rise and show the data used in this study. I perform transdimensional inversion on a synthetic model with both basalts and sedimentary layers to show the effectiveness of the inversion technique. I show the calibration of impedance results with well log data and cores obtained at IODP drill cores on Shatsky Rise. Finally I show the 2D inverted impedance and uncertainty, characterize basalt flows and discuss the implications for late-stage volcanism on TAMU and Ori Massifs.

4.2 Shatsky Rise: Geologic background and data used

Shatsky Rise oceanic plateau is located in the Pacific Ocean ∼1600 km east of Japan (Figure 4.1). Shatsky Rise is formed at a triple junction during late Jurassic and the duration for the formation of the entire plateau is believed to be around ∼21

million years (Zhang et al., 2012). However, the individual volcanoes are formed rapidly (∼1.2 million years) (Sager and Han, 1993). The Shatsky Rise oceanic plateau consists of three volcanoes, the TAMU Massif, Ori Massif and the Shirsov Massif. The TAMU Massif is the oldest of the three volcanoes and Shirsov Massif is the earliest. The oceanic plateau formed as a result of initial high rates of lava effusion during the late Jurassic resulting in the formation of TAMU Massif which is one of largest volcanoes on Earth. The Ori and Shirsov massives are formed at a later stage during a phase of waning effusion rate of lava flows. Massive sheet like basalts are formed due to large pulses of lava effusion and are widely spread often continuous from the summit to the flanks (Zhang et al., 2012). Pillow basalts are formed due to localized explosive volcanism and are often intermingled with volcanoclastic sediments as observed in drilling cores. Shatsky Rise is identified as one of its kind among all the oceanic plateaus on Earth because of its large size and it remained subaqueous since its inception. Almost the entire plateau was below sea level since its formation except the summit of the Ori Massif which is exposed to subaerial conditions evident from the paleosol identified in the core drilled near the summit of Ori Massif.

4.2.1 Data used

2-D MCS reflection profiles are collected by R/V *Marcus G. Langseth* on TAMU and Ori massives as part of the IODP Expedition 324 in 2012 (Thick blue lines in Figure 4.1). The seismic lines are acquired at a shot spacing of 50 m and the nominal fold is 59. The processing steps to obtain the post-stack seismic data include geometry setup, trace edit, band-pass filter, deconvolution, velocity analysis, normal move out and stacking (Zhang et al., 2012). 2D MCS data shows the detailed structure of the basement and also intra basement reflectors compared to an earlier survey obtained in 2010. The seismic data shows reflections upto ∼2 s of two-way travel time which

corresponds to ∼2-4 km depending on the thickness of sediment cover and assuming a velocity of ∼4 km/sec for shallow igneous basement. Line 1 and Line C collected across TAMU and Ori Massifs (Figures 4.2 and 4.3) are used in this study. The data within the rectangles on Line 1 and Line C highlighted in Figures 4.2 and 4.3 is used for inversion. The motivation is to choose seismic traces from summit to the flank of TAMU and Ori Massifs to perform inversion. Several secondary cones are identified on the flanks of both TAMU and Ori Massifs. The secondary cones are formed as a result of uplift of basement rocks due to tectonic activity near the flanks of volcanoes or due to accumulation of eroded volcanoclastics.

In addition to the 2-D seismic reflection profiles, drilling cores ranging from 80 - 250 m are recovered at five locations U1346, U1347, U1348, U1349 and U1350 during the IODP 324 expedition. U1346 is drilled on the summit of Shirsov Massif; U1347 and U1348 are drilled at the east and north flanks of the TAMU Massif; U1349 and U1350 are drilled on the summit and flank of Ori Massif respectively. Locations of U1347, U1349 and U1350 are shown on seismic lines Line 1 and Line C (Figures 4.2 and 4.3). The percentages of basalt types measured from the cores are used for calibration of the inversion results. Cores at U1346 and U1348 are not used in this study due to absence of seismic data on Shirsov Massif and U1348 is drilled on a secondary cone and does not reach the basement. Apart from 2-D MCS lines and drill cores, sonic and density logs are measured in a 170 m interval at U1347 located on the flank of TAMU Massif. The log measurements are used for calibration of seismic inversion results and measure impedance cutoff values for different lithologies which are used for characterizing basalt lithologies from 2-D inverted impedance results.

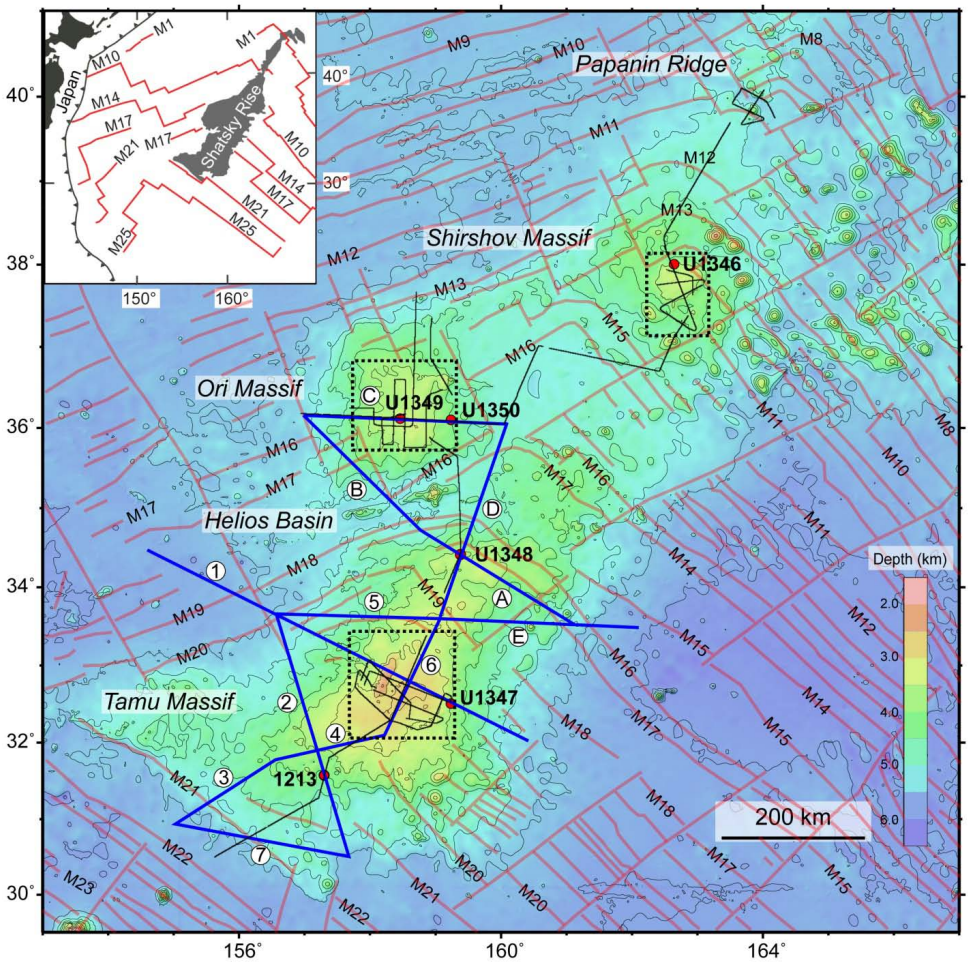


Figure 4.1: Location of Shatsky Rise from Japan shown in inset. The bathymetry map is a satellite predicted map with 500 m contours for depth. Line 1 and Line C across TAMU and Ori Massifs are the seismic data used in this study shown in Figures 4.2 and 4.3. The red dots are the location of ODP and IODP drilling cores. (Figure from Zhang et al 2012)

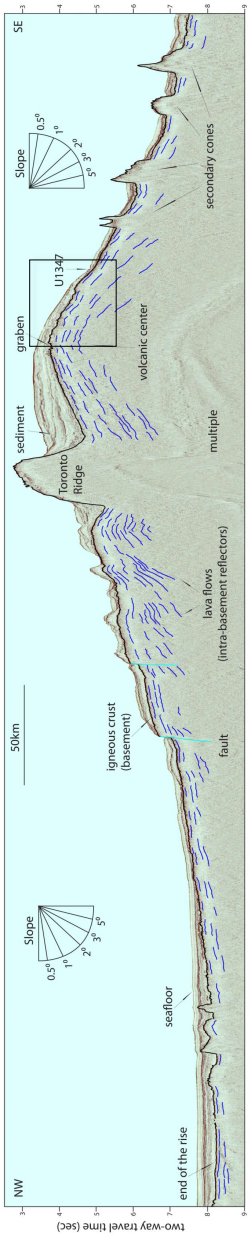


Figure 4.2: Seismic reflection profile of Line 1 shown in Figure 4.1 across TAMU Massif. The thick black horizon is the top of igneous basement. The blue lines are intra basalt reflections. Rectangle section shows the area used to run impedance inversion. The location of IODP drill site U1347 is pointed with an arrow. (Figure from Zhang et al. (2012))

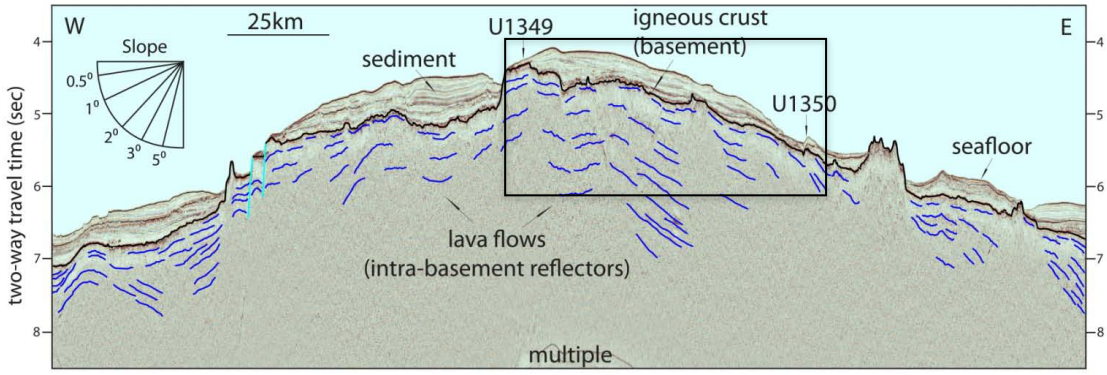


Figure 4.3: Seismic reflection profile of Line C shown in Figure 4.1 across Ori Massif. Rectangle section shows the area used to run impedance inversion. IODP drill site locations U1349 and U1350 at the summit and flank of the Ori Massif are pointed. (Figure from Zhang et al. (2012))

4.3 Method

Post-stack seismic inversion is used to invert for impedance from seismic data. Impedance is a rock property and can be used as a lithology discriminator. The propagator matrix method (Gibson 2005) which generates both primaries and internal multiples is used for forward modeling. The assumptions while using propagator matrix include horizontal layers with a plane wave propagation. As shown on the seismic lines the slopes of the TAMU and Ori massives are less than 2% and hence horizontal layers is a reasonable assumption. The non-linear nature of the forward model would mean using deterministic inversion approach can trap the sampler in a local minimum. To avoid trapping of the solution at a local minimum, I chose to use a stochastic inversion approach in a Bayesian framework. Optimization techniques like simulated annealing are often used to search a large parameter space and force the sampler to find a global minimum. The optimization techniques are useful for finding

the modes of the posterior distribution, however, ensemble of models comprising the posterior distribution does not represent the full posterior distribution. In this study, no optimization technique is used i.e. the likelihood function is a true likelihood function and is not penalized. The reversible jump MCMC sampling technique is used to sample the model space by allowing the number of layers, layer boundaries and impedance to change.

4.3.1 Prior distribution

The initial velocity model and the prior distribution for inversion on Shatsky Rise data are based on velocity obtained from seismic tomography study done by Korenaga and Sager (2012). The low frequency crustal velocity model shown in Figure 9 of Korenaga and Sager (2012) is modified and used as the initial seed model. Korenaga and Sager (2012) applied a generalized linear inversion (GLI) technique to obtain crustal velocity using ray tracing based on refraction and reflection data (21 OBS surveys on a 420 km long refraction transect across TAMU Massif). The prior is an uniform distribution with the higher and lower limits of the prior distributions as, velocity-depth trend obtained from low frequency crustal velocity $\pm$ ΔV_p, where ΔV_p is chosen based on the expected lithology in the interval.

4.3.2 Likelihood function

The likelihood function is a multivariate Gaussian error function. The mismatch between reference data and model seismograms is calculated using root mean square (RMS) error. The reference data in this study is seismic reflection data collected on TAMU and Ori Massifs. A detailed Bayesian formulation to calculate the acceptance criteria for the Metropolis-Hastings ratio is shown in the Appendix. The data error variance is depth independent and no correlation exists between data points. The error standard deviation is chosen as a percentage of maximum amplitude or average

amplitude of the seismic data.

4.4 Synthetic model test

The synthetic model is a 1 km interval with basalt and sedimentary layers. The synthetic model has 4 layers with maximum and minimum layer thicknesses being 600 m and 50 m respectively. The minimum number of layers required to invert the synthetic model will be 20 (Interval thickness/Minimum layer thickness) when using a fixed dimensional stochastic inversion approach. When there is a large variation in layer thicknesses, the transdimensional inversion approach, which allows the number of layers and layer boundaries to change in additional to perturbation in impedance is computationally more efficient than the traditional fixed dimensional inversion.

Figure 4.4 shows the velocity structure of the synthetic model and schematics for primary and internal multiples. The velocity of water is 1500 m/sec and that of shallow sea sediments is 2000 m/sec. The model has 4 layers with shallow sea sediments, basalt layers and sedimentary rocks. The reference seismogram for this synthetic study is generated using propagator matrix method (Gibson 2005) using a source frequency of 30 Hz (Figure 4.5). The initial velocity is 4 km/sec in 10 layers and the upper and lower limits of the uniform prior distribution are shown in Figure 4.6. The model seismogram of the initial seed model is shown in Figure 4.5. The number of layers, layer boundaries and velocity values are allowed to change during the random sweep. Figure 4.6 also shows the best velocity model along with an ensemble of 10000 best models. The boundary between the shallow sea sediments and the basalts is identified exactly. The uncertainty of the boundaries and velocity of the layers can be identified from the ensembles of velocity profiles. In some of the ensembles, the absolute impedance values in the ensemble of models is different from the reference values, however, the contrast between impedances is close. Figure

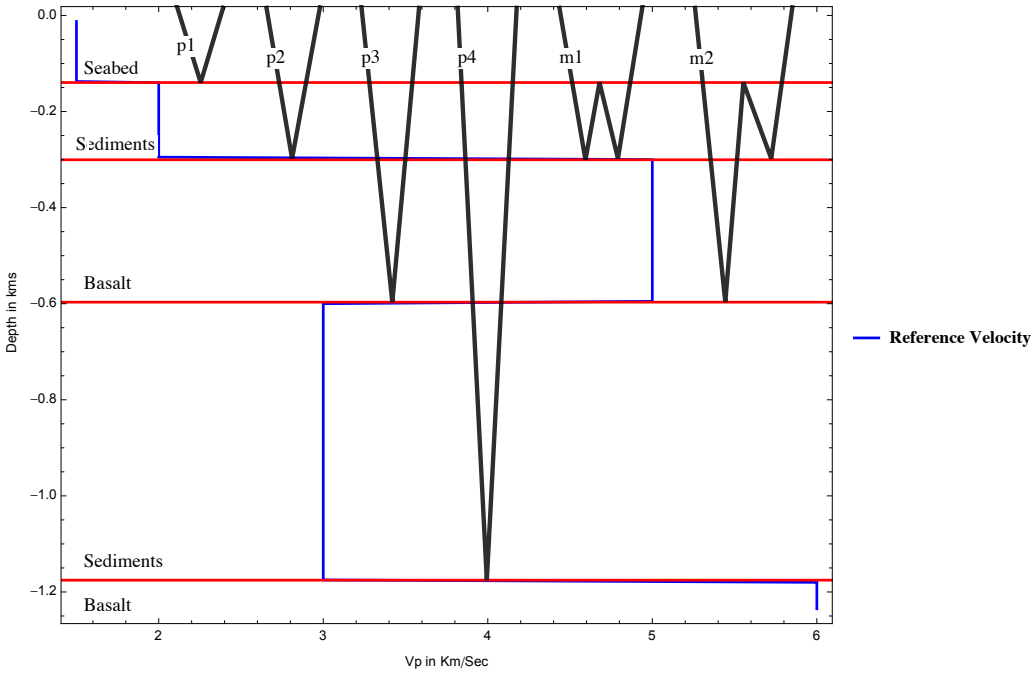


Figure 4.4: Reference velocity profile and schematic of primary and internal multiple reflections.

4.7 shows the histogram of layer boundaries, mean, lower and upper quantiles of the velocities of 10000 best models. The uncertainty in the boundaries and velocities is due to the non uniqueness of solutions. Figure 4.8 compares the reference seismogram with the best model seismogram. Figure 4.9 shows the histogram of number of layers and the variation of number of layers with iterations. The initial model has 10 layers and the histogram has a peak value at 5 layers compared to 4 layers in the reference model.

4.5 Calibration study

In this section, I compare the inversion results with, velocity log at U1347 on TAMU Massif and, lithology from three drill cores on TAMU and Ori massives to validate the inversion results. The three main lithologies identified in the Shatsky

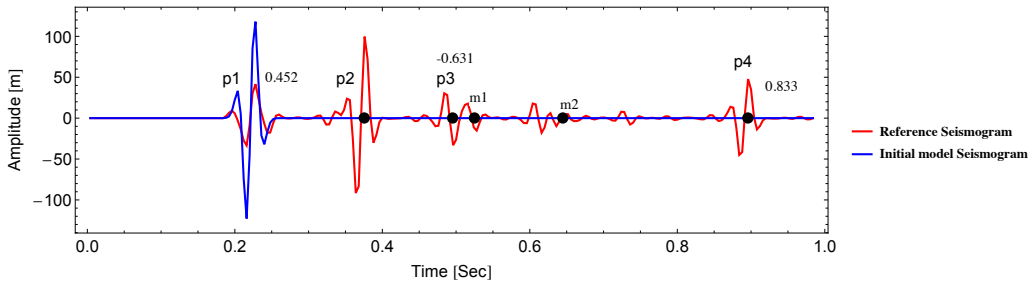


Figure 4.5: (a) Reference seismogram generated from the reference velocity model defined in Figure 4.4 using a source wavelet of 30 Hz using the propagator matrix method. The travel times for primaries (p1, p2 and p3) and two internal multiples (m1, m2) are identified as solid dots on the time axis. The ratio of reflectivity's for p1, p2 and p4 normalized with respect to p2 are also shown.

Rise oceanic plateau include massive basalt flows, pillow basalts and volcanoclastics. Pillow basalts have lower impedance values compared to massive basalts due to the presence of intercalated volcanoclastic layers and high porosity pillow flows in the pillow basalt packages (Zhang et al., 2012; Bartetzko et al., 2002). Based on the inspection of lithologies in the IODP cores, the predominant lithology of TAMU Massif is identified as massive sheet like basalts and that of Shirsov Massif is pillow basalts. The lithology of the Ori Massif located in between TAMU and Shirsov Massif has intermediate proportions of massive basalt flows. The basalt section of the drilled core at U1347 on TAMU Massif has ∼33% of pillow basalts compared to ∼86% at U1350 located on the flank of Ori Massif (Sager et al., 2011). Percentage of pillow basalts in all the cores is shown in Table 4.1 along with their location on the Volcanoes. Figure 4.10 shows the lithology description of the drill cores.

The well log interval and drill cores are 80-160 m thick in the igneous basement. The shallow impedance (∼0.8 km) obtained using 0.4 s (two-way travel time) of post-stack seismic trace at U1347 is compared with the impedance from the well log

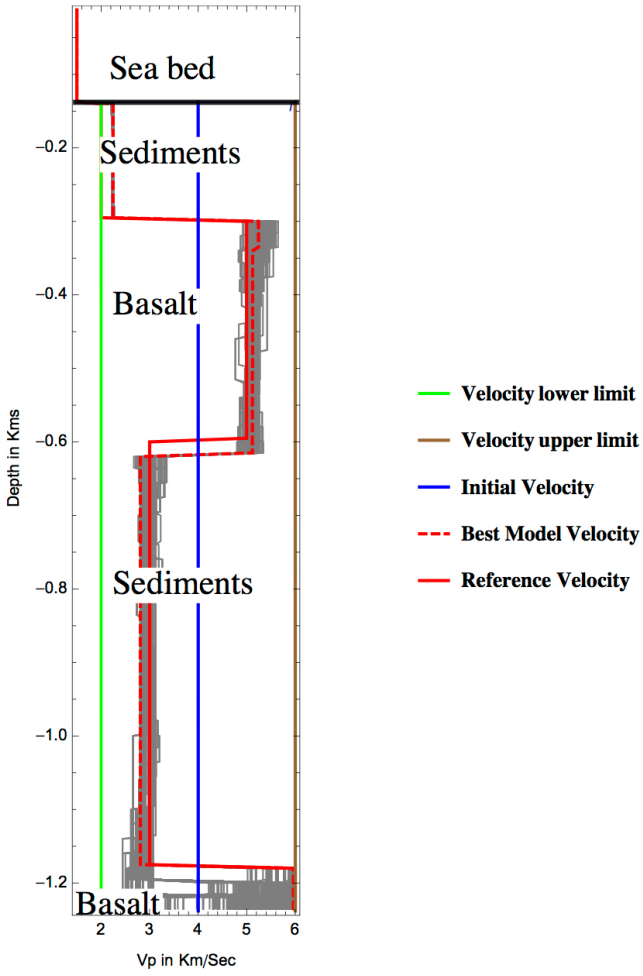


Figure 4.6: Synthetic model used for seismic inversion. The reference velocity model includes stacked basalt flows and sedimentary layers. The green and brown solid lines are the lower and upper limits of prior uniform distributions.

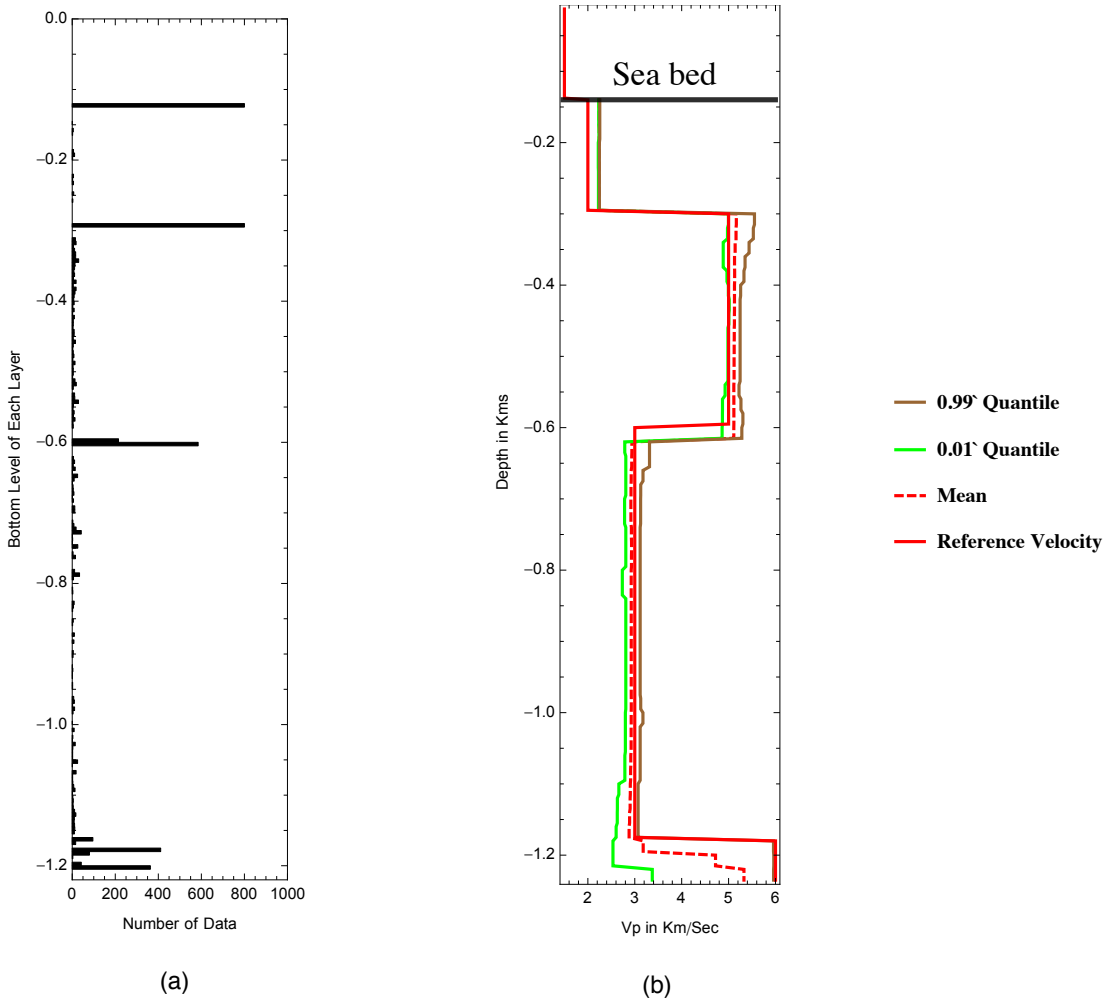


Figure 4.7: (a) Histogram of layer boundaries, after running seismic inversion using RJMCMC sampling. (b) The mean, lower and upper quantiles of inverted velocity models. Both the plots are obtained from 10000 best models out of 50000 models.

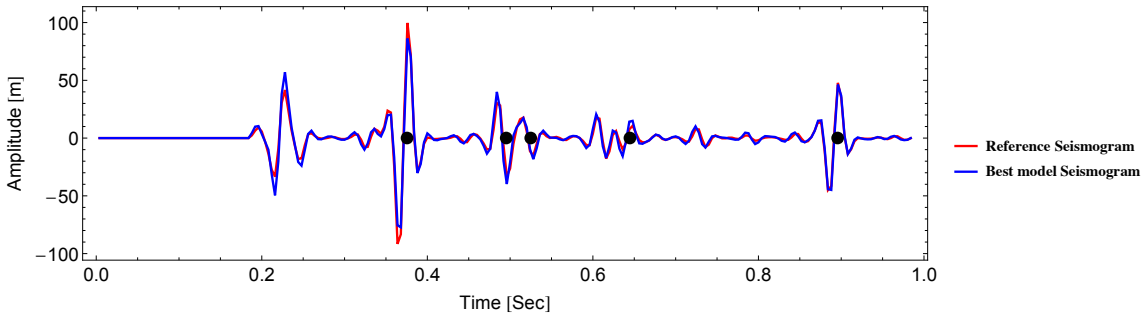


Figure 4.8: Comparison of the reference seismogram and the seismogram generated from the best velocity model.

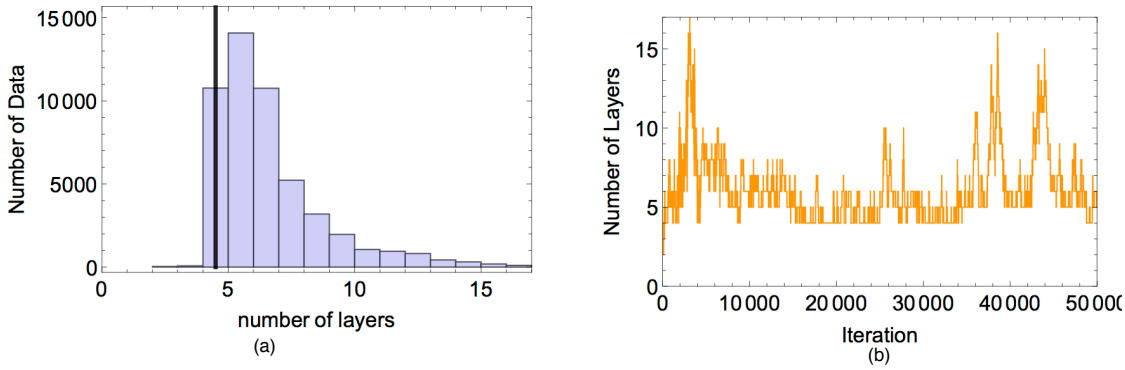


Figure 4.9: (a) Number of layers histogram for all the 50000 iterations. The actual number of layers of the reference velocity model is 4 shown as a thick line on the histogram. (b) Number of layers versus iteration.

measurements at U1347. The amplitude spectrum extracted from seismic trace at U1347 is shown in figure 4.11. A 30 Hz central frequency Ricker wavelet is used for forward modeling using propagator matrix. Figure 4.11 compares amplitude spectrum from seismic trace at U1347 with amplitude spectrum of a 30 Hz Ricker wavelet. The blue solid line in Figure 4.12 is the prior velocity model from Korenaga and Sager (2012). The upper and lower limits of the prior distribution are velocity-depth trend obtained from Korenaga and Sager (2012) $\pm$ 2 km/sec respectively. The range of the uniform prior distribution is chosen to allow all three possible lithologies during inversion. Figure 4.12 compares the inverted velocity at U1347 with the velocity log. The inverted velocity at U1347 generally follows the same velocity trend as the velocity log except for the last basalt interval in the velocity log. A linear relationship between density and velocity obtained from sonic and density logs at U1347 is used in the seismic inversion (Figure 4.13). During the impedance inversion, only independent parameter is velocity and density is estimated from velocity. Interbedded volcanoclastics with thicknesses <10 ft are not identified in inverted velocity results since a higher frequency than the seismic band-pass are required to invert for thicknesses less than 40 m ($\sim\lambda/4$). Zhang et al. (2012) compared the synthetic seismogram obtained from well logs with seismic trace at U1347 and suggested the intra basalt reflectors occur at the contrast of impedance between massive and pillow basalts and between basalts and intercalated sediments. Figure 4.14 compares the seismic trace at U1347 with initial, best model seismograms and an ensemble of the best 20000 seismograms. The velocity model from Korenaga and Sager (2012) is modified to match the thickness of the sediments above the basement to generate the initial velocity model.

The shaded region in Figures 4.12, 4.15 and 4.17 shows the depth interval where drill cores are obtained in igneous basement. The lithology in the cores can be

identified in the order of cm's, however the vertical resolution of seismic inversion results is in the order of 10's of meters. Impedance is a rock property and can be used as a lithology discriminator based on lithology cut offs obtained from well log measurements. Zhang et al. (2012) identified the low velocity interval at 3.75 km on the well log at U1347 (Figure 4.12) as pillow basalts after tying well log and the lithology from core. Lithology can be reconstructed from the physical measurements obtained from well logs. The well log measurements and drill core can be tied to identify the well log measurements corresponding to a given lithology. The range of measurements for each lithology can be used to determine lithology at locations where well logs are not available. Another way of reconstructing lithology is by determining electrofacies from all the available logs using neural networks. This method consists of training a subset of logs measurements to obtain lithology. The later of the two methods is more efficient due to poor recovery of cores and difficulty in lithology correlation based on direct comparison of cores and logs. Bartetzko et al. (2002) reconstructed the lithology from physical measurements using neural networks from boreholes 504B and 896A drilled about 200 km south of the Costa Rica Rift. Hole 504B is the deepest well drilled in oceanic crust and consists of large sections of massive flood and pillow basalts. The average of the velocity and density log measurements for pillow basalts from 504B is 4.7 km/sec and 2.5 gm/cc respectively. The average of the velocity and density log measurements for massive flood basalts from 504B is 5.6 km/sec and 2.7 gm/cc respectively. The ranges of velocity measurements for massive flood and pillow basalts are 3-5.5 and 5-6.5 km/sec in Holes 504B and 896A (Bartetzko et al., 2002). On the basis of results from Bartetzko et al. (2002), impedance values from 8-11.75 km/sec.gm/cc are identified as pillow basalts and values higher than 11.75 km/sec.gm/cc are identified as massive flood basalts. The upper limit of 11.75 km/sec.gm/cc for pillow basalts corresponds to the average

p-wave velocity of 4.7 km/sec from well log observations in Bartetzko et al. (2002). I compare the percentage of lithologies from drill cores with those inferred from inverted impedance (Table 4.2).

The prior velocity from Korenaga and Sager (2012) is available only across TAMU Massif since the OBS survey are collected only on TAMU Massif. The prior for Bayesian inversion at Ori Massif is obtained by modifying the shallow sediment thickness in the velocity-depth trend from TAMU Massif. Figures 4.16 and 4.18 compare the seismic traces at U1350 and U1349 on Ori Massif with initial and best model seismograms. Percentage of basalt types obtained from cores and inverted impedance values are in agreement at U1347 and U1350. Lithology description of core at U1349 show only massive basalt flows, however, some of the impedance values from inversion are lower than 11.75 km/sec.gm/cc suggesting presence of pillow basalts. The low impedance values of basalt at U1349 is due to the high level of geochemical alteration of basalts at the summit of Ori Massif due to subaerial exposure and shallow water lava flows (Sager et al., 2011, 2013; Delacour and Guillaume, 2013). Table 4.2 shows the level of alteration at each of the drill core sites and the percentage of impedance values below the cut off defined for pillow basalts. The results at U1349 suggest impedance values lower than that of massive flow basalts can be due to presence of pillow basalts or alteration of basalts caused by to fluid-rock interactions.

4.6 Inversion results

The seismic profiles Line 1 and Line C on TAMU and Ori Massif respectively are too large to perform transdimensional inversion for the complete profiles. Hence, I selected portions of the seismic profiles on the shallow structures of TAMU and Ori massifs which include the strong reflection at the basement and few intra basement reflections. The objective is to calculate acoustic impedance from post-stack seismic

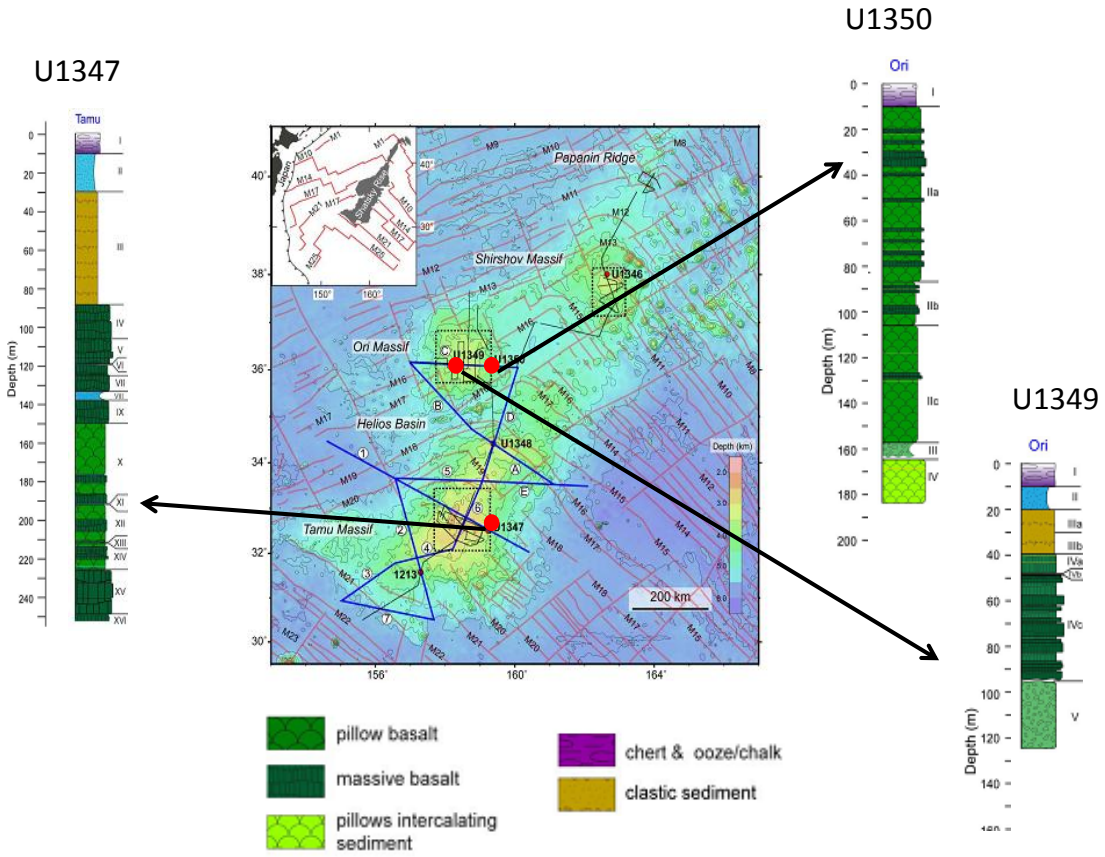


Figure 4.10: Location of IODP drill sites U1347, U1349 and U1350 shown as red dots on the Shatsky Rise bathymetry map and lithology description of the drill cores.

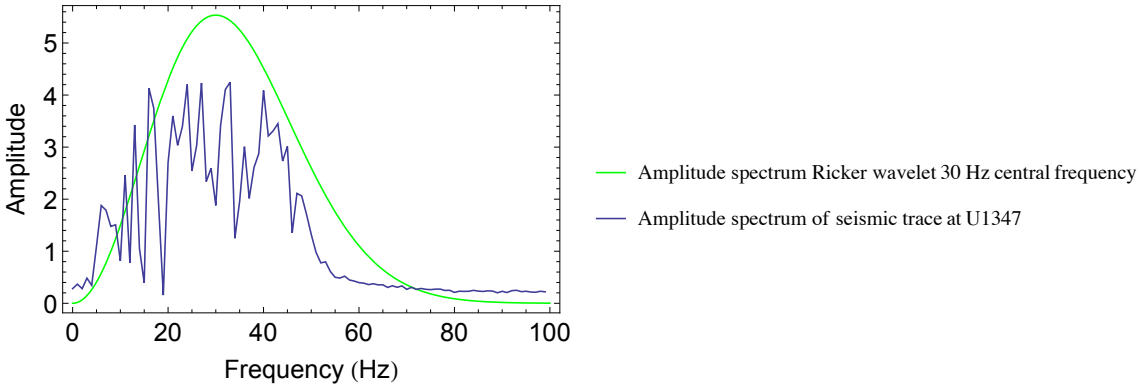


Figure 4.11: Amplitude spectrum of 1 sec of seismic trace at U1347.

Table 4.1: Table showing the percentage of pillow basalts and degree of alteration at the IODP drill sites.

Well Name	Location	% of pillow basalts from drilled core (Sager et al 2012)	Degree of alteration (Sager et al 2012, Delacour and Guillaume 2013)
U1347	TAMU Massif flank	~33	Light low-temperature fluid-rock interaction
U1349	Ori Massif summit	0	Highly altered (subaerially exposed and lava flows may be deposited in shallow water)
U1350	Ori Massif flank	~86	Slightly to moderately altered submarine lava flows

Table 4.2: Table comparing percentage of basalt types in drill cores with those inferred from inversion results. The inversion results are obtained in the same interval as the drill core interval.

Well Name	Location	% of pillow basalts from drilled core (Sager et al 2012)	Degree of alteration	% of impedance values below the cut off value of pillow basalts
U1347	TAMU Massif flank	~33	Light low-temperature fluid-rock interaction	35.7
U1349	Ori Massif summit	0	Highly altered (subaerially exposed and lava flows may be deposited in shallow water)	100
U1350	Ori Massif flank	~86	Slightly to moderately altered submarine lava flows	75

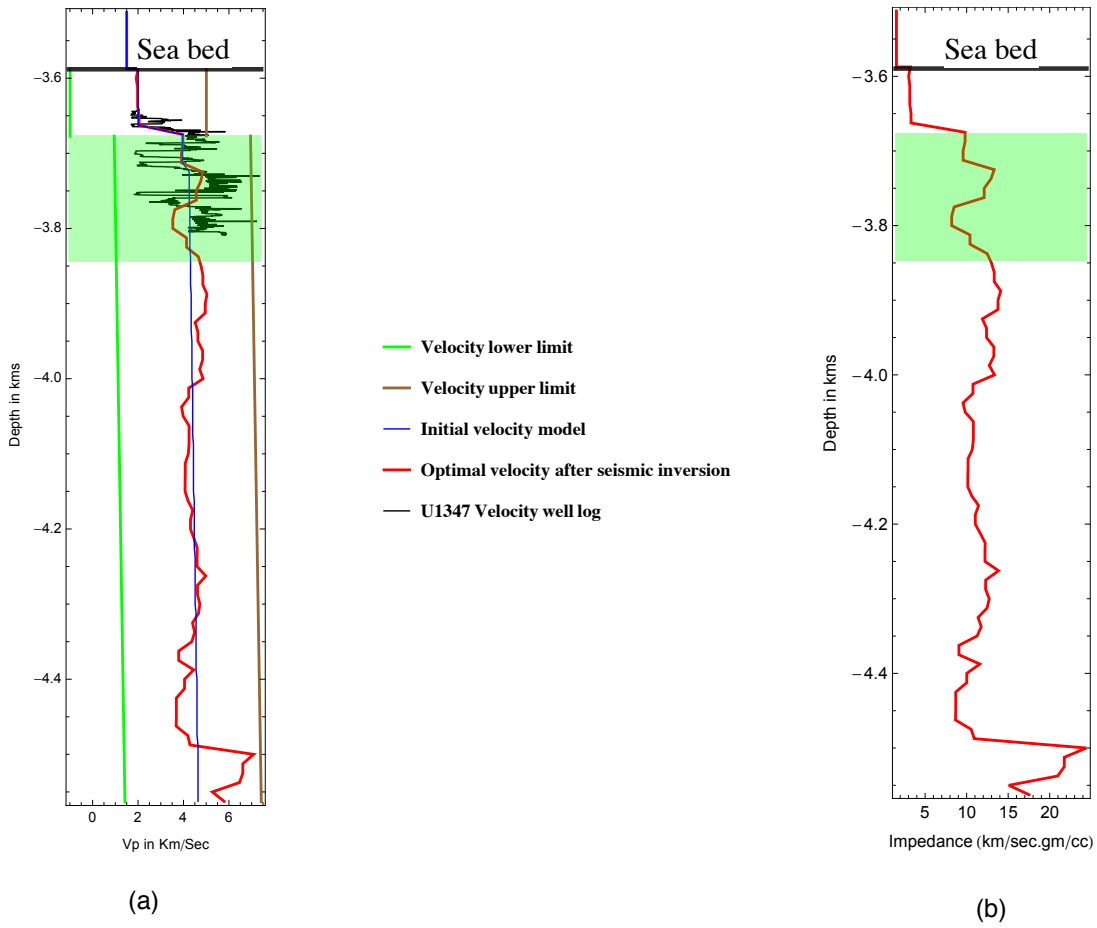


Figure 4.12: (a) Best model velocity from inversion using post-stack seismic data, at location U1347 on TAMU Massif, as reference data compared with the velocity log obtained at U1347. The well log is measured only for 170 m interval as shown in the figure. The shaded interval is where a drill core is collected by IODP expedition 324. (b) Best model impedance.

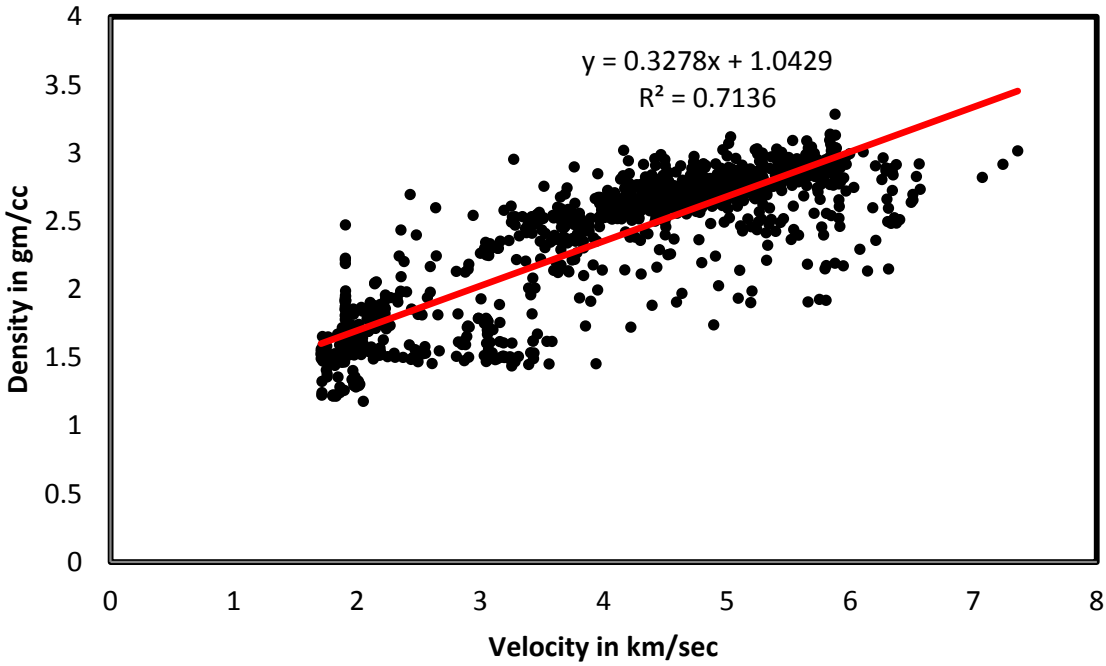


Figure 4.13: Velocity and density relationship obtained from the well logs at U1347 on TAMU Massif.

data, discriminate lithologies based on impedance results and compare the lateral and vertical extent of massive flow basalts and pillow/altered basalts on the two seismic lines. The location of seismic data subsets picked for inversion are shown in rectangles on the complete seismic profiles in Figure 4.2 and 4.3. The subsets of data are chosen to include seismic traces from summit to the flank of the volcanoes.

The reversible jump MCMC algorithm used for inversion randomly chooses one of three moves during the random sweep. The three moves include perturbation of a layer boundary and velocity in a layer simultaneously, adding a layer and perturbation of velocity in a layer simultaneously and third type of move is deleting a layer and simultaneously perturbing velocity in a layer. Impedance inversion is obtained for a single seismic trace at a time which implies the inversion is a 1-D inversion. However, the results from all the 1-D models are put together for visualization and interpretation

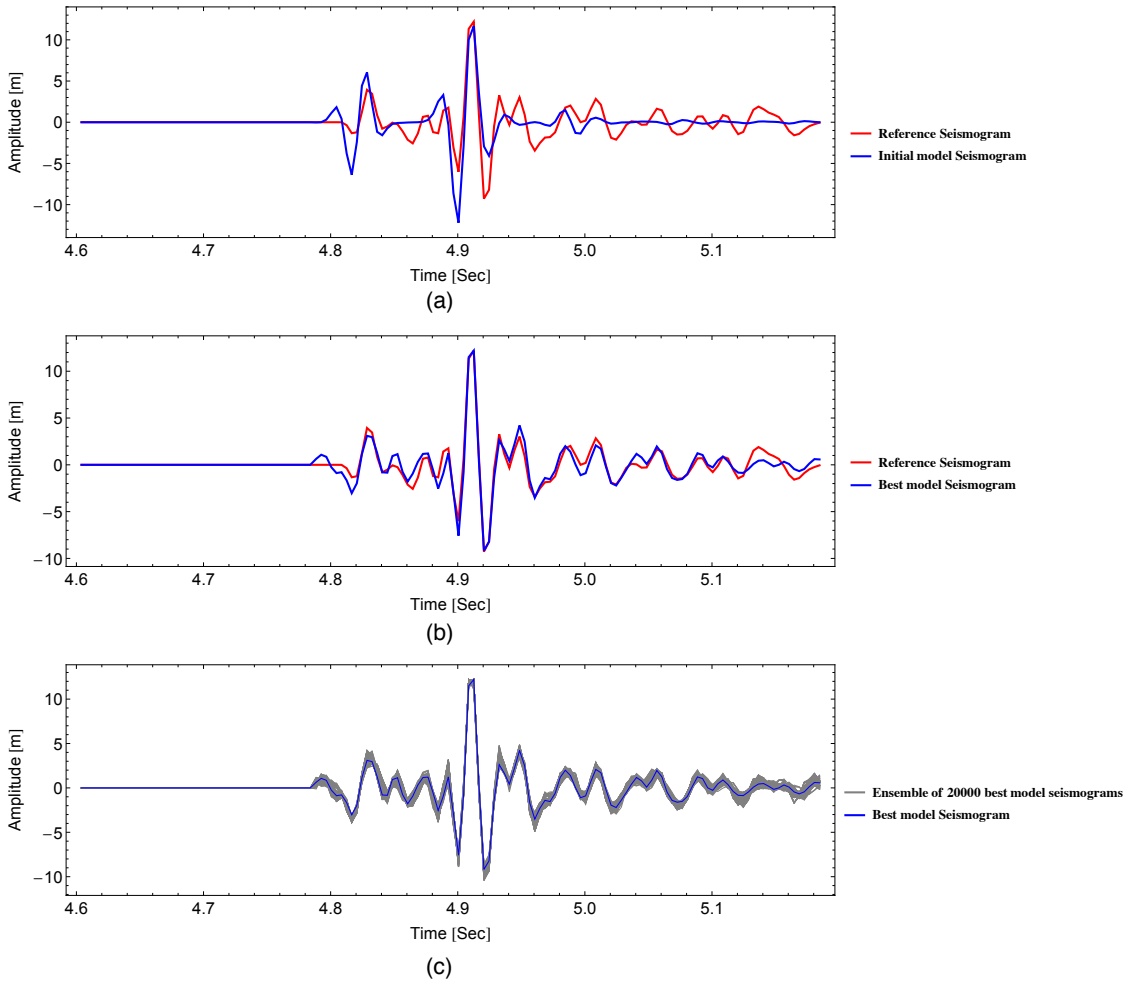


Figure 4.14: Seismic trace at U1347 compared with (a) initial model seismogram and (b) best model seismogram. (c) Ensemble of 20000 best model seismograms.

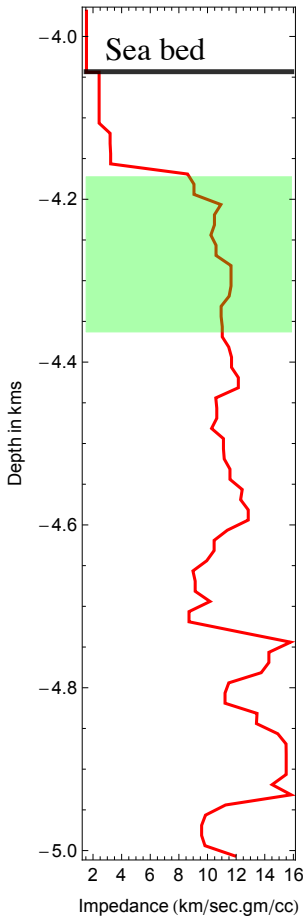


Figure 4.15: Impedance result after post-stack seismic inversion at location U1350 on Ori Massif. The shaded interval is where a drill core is collected by IODP expedition 324.

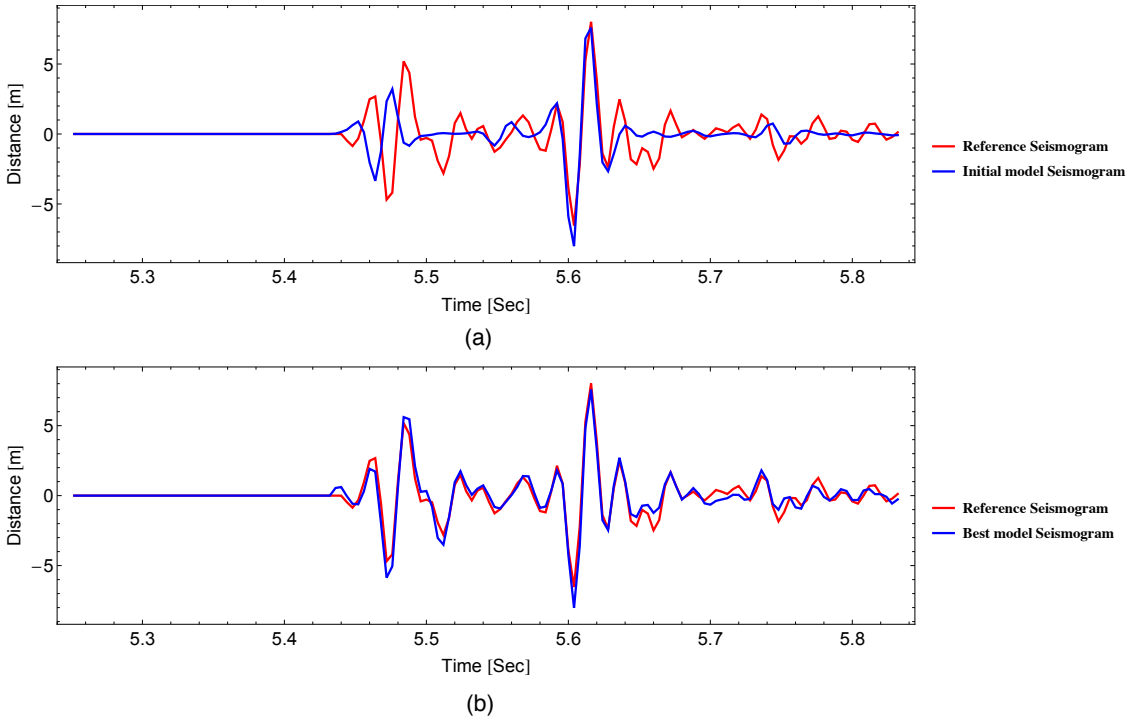


Figure 4.16: Seismic trace at U1350 compared with (a) initial model seismogram and (b) best model seismogram.

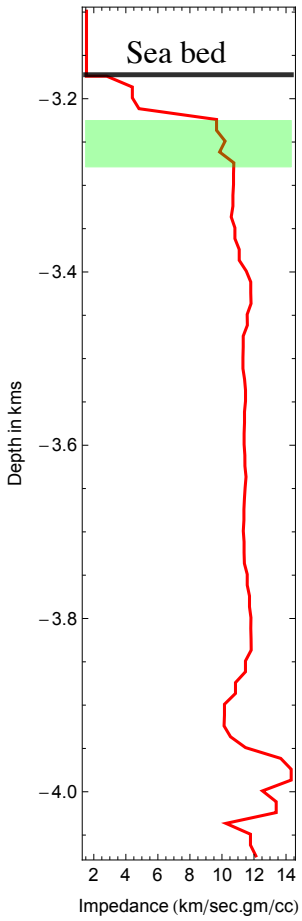


Figure 4.17: Impedance result after post-stack seismic inversion at location U1349 on Ori Massif. The shaded interval is where a drill core is collected by IODP expedition 324.

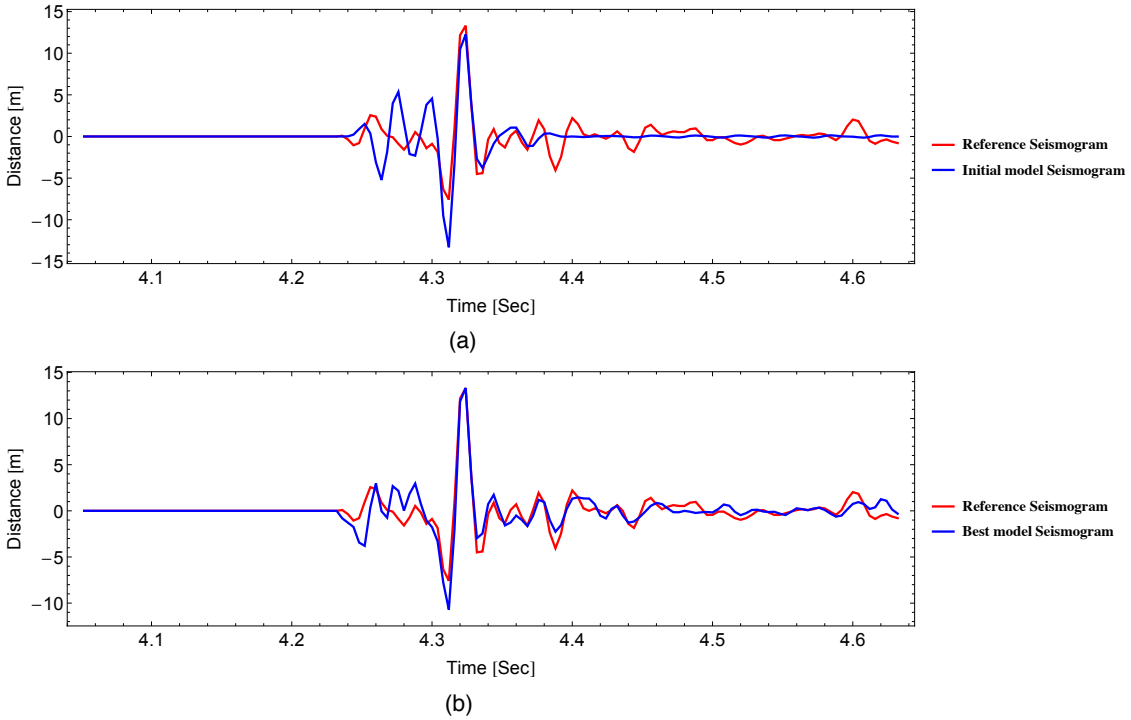


Figure 4.18: Seismic trace at U1349 compared with (a) initial model seismogram and (b) best model seismogram..

purposes. The error function is a multivariate Gaussian distribution with a standard deviation of 0.05 for all the data points which implies approximately ±2% of error is allowed for acceptance. Standard deviation of the proposal distribution which defines the jump of model parameters at each iteration is 1. The standard deviation of the error function and proposal distribution, and the coefficient of walk size which defines the amount of movement of layer boundary at each iteration, are remained as constant for all the seismic traces. The forward model is generated using propagator matrix method and a Ricker wavelet with a central frequency of 30 Hz for source frequency.

4.6.1 Line 1 on TAMU Massif

The complete seismic profile of Line 1 on TAMU Massif is ∼700 km, but the analysis in this chapter is focused on 52 km by 1 sec section of data from summit to the flank of the TAMU Massif. The thick black solid line in Figure 4.2 is the top of the igneous basement and can be traced all along the seismic profile. The blue lines in the figure are referred to as intra basalt reflections by Zhang et al. (2012) are caused by difference in impedance between basalt types or between basalts and volcanoclastics. When a seismic trace is chosen for every 1 km from summit to the flank of TAMU Massif, the total number of seismic traces are 52 which are shown in Figure 4.19.

The prior model is chosen based on the velocity profile obtained from Korenaga and Sager (2012) seismic tomography study. The low frequency velocity profiles obtained from Korenaga and Sager (2012) are available only at every 10 kms. For inversion of seismic traces in between the locations where prior velocity models are available, I chose the nearest prior velocity model and adjusted the sediment thickness values to prepare a prior model which can generate a good match near the reflection

at the top of the igneous basement (Figure 4.20). Figure 4.19 compares the seismic reflection data with the initial model seismogram. There is a thicker sediment cover at the summit compared to that at the flank of the TAMU Massif. Since the length of seismic trace (1 s of TWT) used for inversion is constant for all the traces from summit to the flank, the thick sediment cover at the summit implies the interval of igneous basement where impedance inversion is performed is lower than that at the flanks. Sager et al. (2013) interpreted the graben at the summit of the TAMU Massif as the caldera of the volcano since most of the intra basalt reflection diverge and extend to the flank of the volcano. The individual thickness of massive flow basalts identified in the IODP drill core U1347 at TAMU Massif are upto ∼23 m. Since the vertical resolution of the seismic data is ∼44 m, only basalt flow packages which are stacked basalt types can be identified from the inversion results.

The 1-D best impedance models at each of the seismic trace locations are put together and interpolated to generate a 2-D shallow impedance profile on the TAMU Massif (Figure 4.21). One of the advantages of following a stochastic inversion approach include quantifying uncertainty in impedance. Figure 4.22 shows the standard deviation of impedance. The standard deviation is high near 18 km as the seismic trace is a bad trace due to processing error. Except for the areas with bad seismic traces a large area in the cross section has standard deviation values between 0.33 to 0.66 km/sec.gm/cc. Figure 4.23 compares the seismic data with the best model seismograms for ∼1 sec of seismic data. A lithology profile is generated from the 2-D impedance profile using the lithology discriminator cut offs mentioned in the calibration study. Figure 4.28 shows the lithology distribution comprising of sediments, pillow/altered basalts and massive flow basalts. The heterogeneities in the shallow basement structure both in the vertical and lateral directions are evident in the impedance structure which result in reflections referred to as intra basalt

reflectors by Zhang et al. (2012).

4.6.2 Line C on Ori Massif

The seismic profile on Ori Massif, Line C, is ∼300 km and the seismic section used in this study for inversion is a 80 km by 1 sec area. The seismic section extends U1349 near the summit of the volcano to U1350 at the flank. A seismic trace is extracted at every 2 km which implies a total number of 40 traces are used for inversion. The prior impedance model at Ori Massif is obtained by modifying the velocity profile obtained from Korenaga and Sager (2012) in a similar way as mentioned for TAMU Massif. Figure 4.24 shows the initial impedance model at Ori Massif and Figure 4.25 compares the seismic data with the initial model seismogram. There are two local highs at the igneous basement and a ∼10 km wide graben identified in Figures 4.24 and 4.25. The local highs may be due to uplift of fault blocks, vertical accumulation of eroded volcanoclastics due to lateral constrain or pillow basalts accumulated at a low effusive rate. The graben is close to the summit of the Ori Massif and questions are raised about the graben being the caldera of Ori Massif. However, the symmetry and divergence of reflection to the east and west of the summit are not as convincing as in the case of the caldera at TAMU Massif. The lateral extent of the caldera of TAMU Massif is only 2-3 km and given the size of TAMU Massif is larger the Ori Massif, a caldera at Ori Massif which is twice as big as that of TAMU Massif is unlikely. More seismic coverage near the caldera of Ori Massif can be useful in making inference about the location of caldera of Ori Massif. Figure 4.26 shows the standard deviation of impedance at Ori Massif. The range of standard deviation values at Ori Massif is similar to that at TAMU Massif and lie between 0.33 to 0.66 km/sec.gm/cc. Figure 4.27 compares the seismic data with the best model seismograms for ∼1 sec of seismic data at Ori Massif.

4.7 Discussion

TAMU Massif is formed in the initial stage of the Shatsky Rise oceanic plateau formation when the effusion rate of lava is high and resulted in frequent massive lava flows. The wide spread lava flows are recorded as massive flow basalt rocks in the geologic record. The massive flows waned during the formation of Ori and Shirsov Massifs which is observed from the lithology of the drill cores on Ori and Shirsov Massifs. The IODP and ODP cores at U1347 and 1213 show thick basalt flows with individual layers ranging upto ∼23 m (Sager et al., 2013; Koppers et al., 2010) and higher percentage of massive basalt flows compared to pillow basalts. Massive flow basalts are also identified on Ori and Shirsov masses but the thickness and frequency of their occurrence is lower compared to that at TAMU Massif. The percentage of massive basalt flows in the section on TAMU Massif is higher than that on Ori Massif as observed from Figures 4.28 and 4.30. The comparison of percentage of massive flows basalts from inversion results in the shallow structure (∼1 km of igneous basement) is generally in agreement with comparison made from core data (∼80-160 m) and with the evolution history of TAMU and Ori Massifs.

Impedance for basalts lower than that of massive flow basalts can be due to presence of pillow basalts or geochemically altered basalts. Pillow basalts are formed in an environment of local explosive volcanism and they are porous and interbedded with inter-flow sediments. Alteration of basalts is similar to diagenesis, where the parent rocks can be massive flow or pillow basalts but subsequent geochemical alteration can replace basalt minerals with secondary minerals. X-ray diffraction (XRD) analysis done on rock powder by Delacour and Guillaume (2013) from the two highly altered ($80 - 95\%$ alteration) cores at summits of Ori and Shirsov Massifs suggest the basalt minerals are primarily replaced by clay minerals especially smectites

and calcite. The degree of alteration depends on the fluid-rock interactions at various temperature and redox conditions. The temperature gradient between the volcano and the surrounding sea water is high at the summit compared to the flank which results in a high degree of alteration of the basalts at the summits Delacour and Guillaume (2013). The degree of alteration is a function of structural position on the volcano and also vertically at a single location. Table 4.2 shows the degree of alteration from the three cores studied in this study and their relative locations on the volcanoes.

The percentage of pillow/altered basalts at TAMU summit are higher compared to that at the flank as observed from Figures 4.28. The higher percentage of pillow/altered basalts at the summit can be due to frequent localized explosive volcanism at the summit of the volcanoes forming pillow basalts, or due to high degree of alteration of basalt rocks at the summit. Figure 4.29 shows the variation of pillow basalt percentage with distance for three different pillow basalt impedance cut off values of 10, 11 and 12 km/sec.gm/cc respectively. The decreasing trend in pillow basalts from summit to flank of TAMU Massif is evident irrespective of the variation in cut off values for impedance of pillow basalts. I am showing the variation of pillow basalts with different cut off values to represent the uncertainty in the average impedance values for pillow basalts. Notice the difference in the impedance cut off values (1 km/sec.gm/cc) should be greater than the standard deviation in impedance results ($\sim$0.33-0.66 km/sec.gm/cc) to make a credible inference about the spatial variation in percentage of pillow basalts. Since TAMU Massif stayed subaqueous since its inception, the low impedance of basalts at the summit is mostly likely due to pillow basalts rather than altered basalts. A similar comparison of percentage of pillow basalts at summit and flank is not valid at Ori Massif. This might be due to the fact that the area of interest on Ori Massif is entirely exposed to

subaerial conditions and the lavas flows are deposited in shallow marine environment implying the low impedance is most likely due to alteration (Figure 4.30). Figure 4.31 shows alteration of basalts is identified in the entire area on Ori Massif irrespective of pillow/altered basalt cut off value. The volume of pillow/altered basalts is higher on Ori Massif compared to TAMU Massif which might be due to the waning of massive flood basalts during the formation of Ori Massif. Another possible reason for the difference in percentage of low impedance basalts on Ori Massif compared to TAMU Massif is due to the variation in geochemical composition of source magma types which build the volcanoes. Sano et al. (2012) studied 99 lava basement samples from the IODP core intervals with additional data from Mahoney et al. (2005) to study the composition of magma source for each of the volcanoes and found the TAMU Massif is composed of ∼94% of mid-ocean ridge basalts (MORB) and Ori Massif is composed of ∼57% of non-MORB type basalts. The effect of different magma types on rock properties like impedance is unknown and needs further investigation.

4.8 Conclusions

The 2-D MCS reflection data at Shatsky Rise obtained during the IODP expedition 324 show good penetration into the shallow igneous basement inspite of the large reflection at the basement and shallow sediment interface and due to internal scattering of energy in the basalts. Impedance inversion results from the shallow crust (∼1 km) suggest higher percentage of massive flow basalts at TAMU Massif compared to Ori Massif. The percentage of pillow/altered basalts is higher at the summit compared to that at the flanks of the volcanoes. This might be due to higher level of localized explosive volcanism and/or geochemical alteration at the summit compared to that at the flank of the volcanoes.

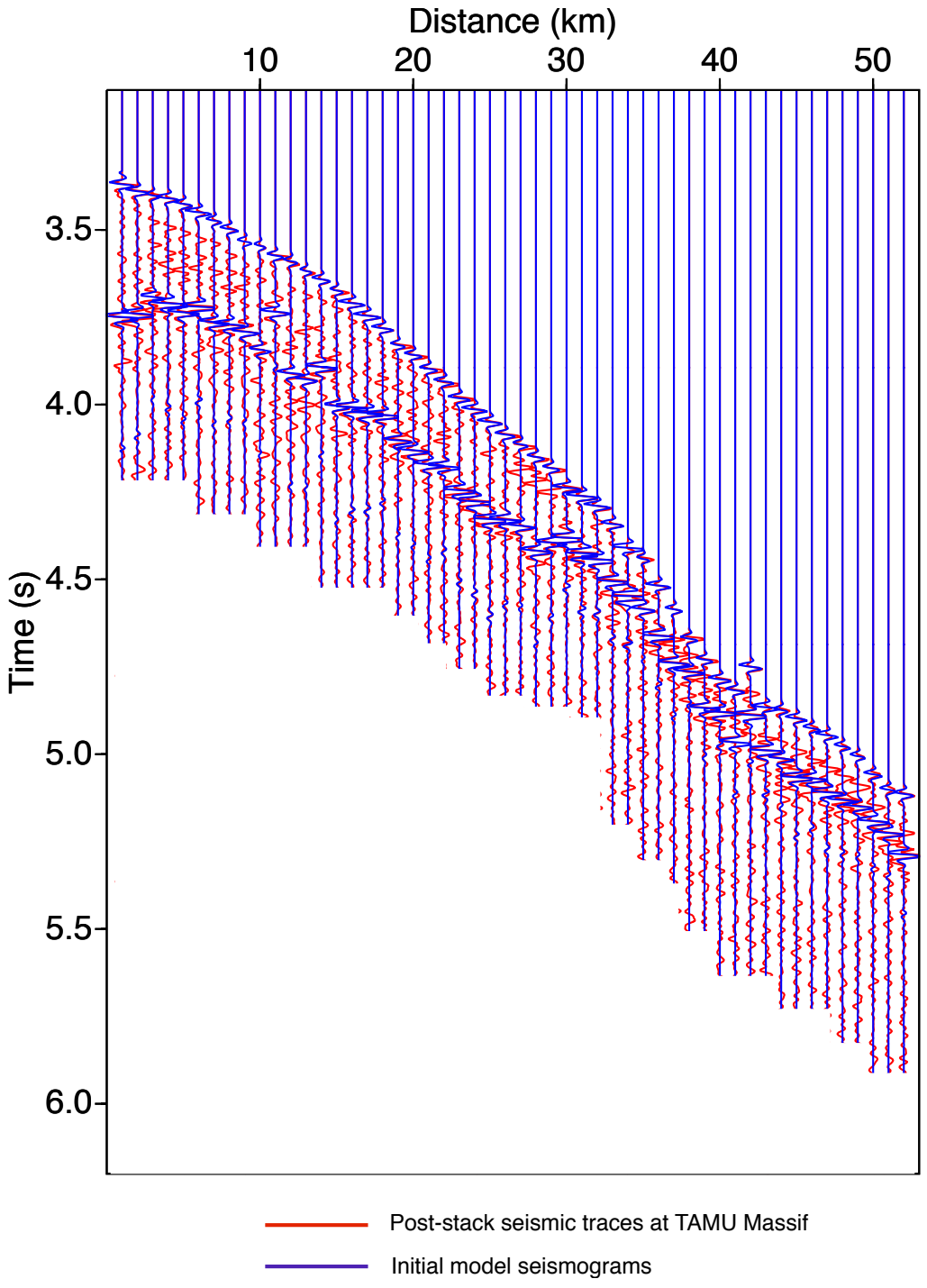


Figure 4.19: Comparison of seismic traces at the area of interest on TAMU Massif with the seismograms generated from prior impedance model. The distance between seismic traces is 1 km. The location of the seismic section chosen here is shown as a rectangle on TAMU Massif in Figure 4.2.

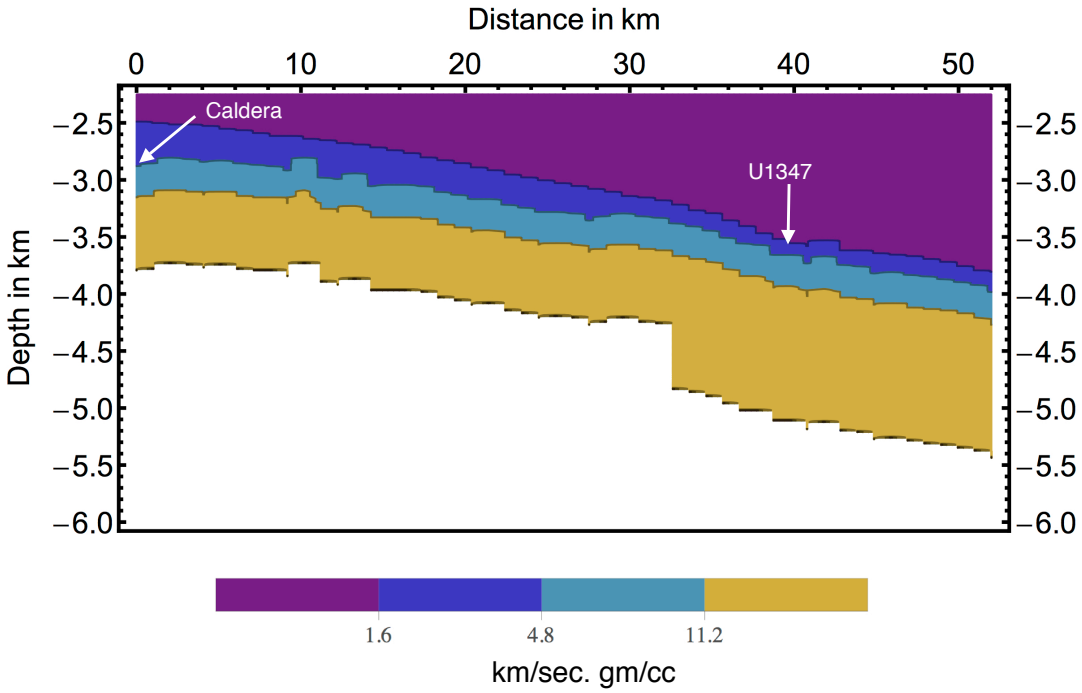


Figure 4.20: The prior impedance model in the area highlighted on TAMU Massif in Figure 4.2. The prior for velocity is obtained from crustal seismic tomography study conducted by Korenaga and Sager (2012).

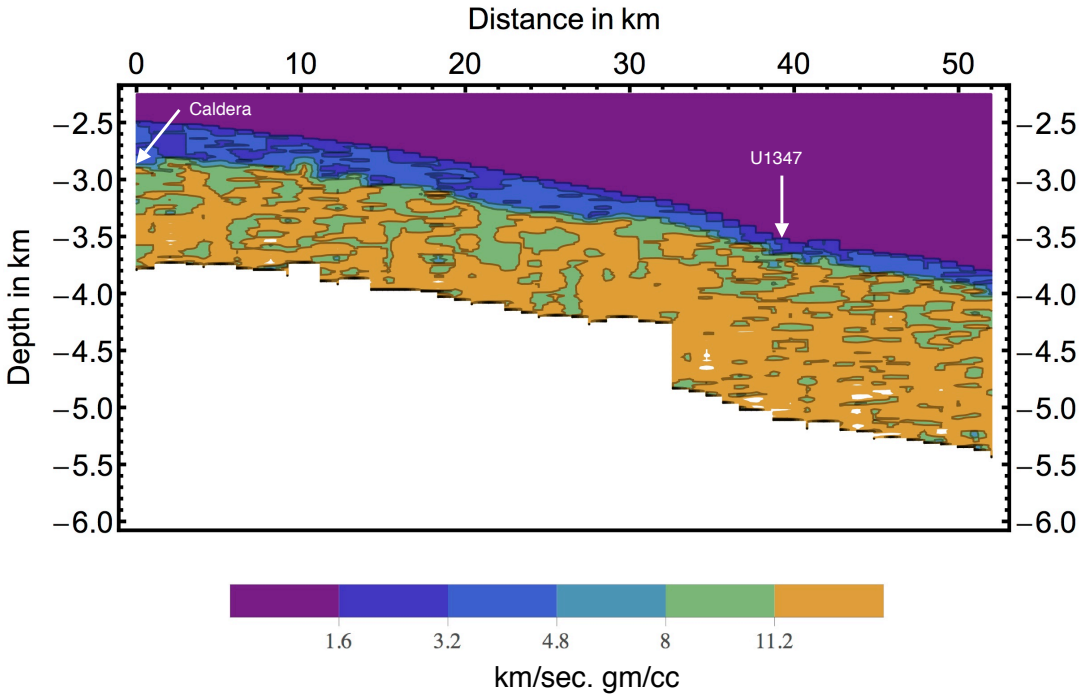


Figure 4.21: Best impedance model in the area of interest on the TAMU Massif.

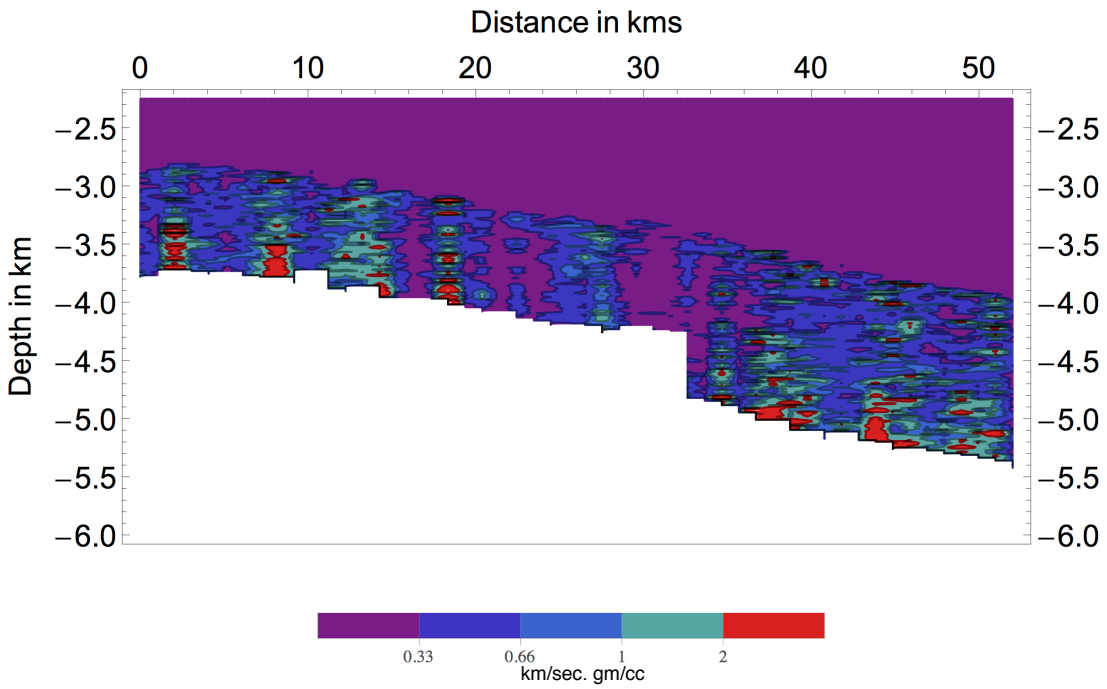


Figure 4.22: Standard deviation of impedance models in the area of interest on the TAMU Massif.

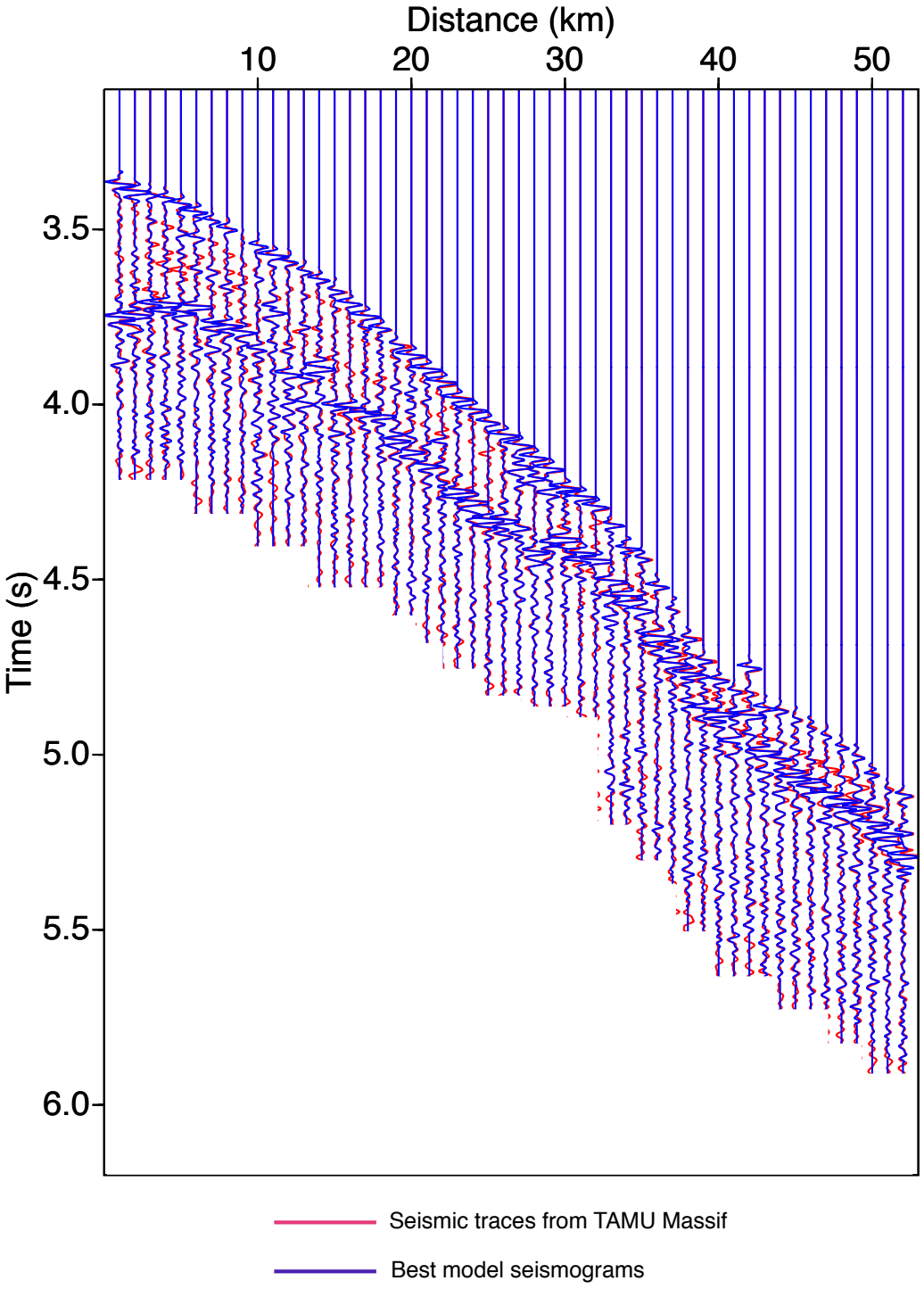


Figure 4.23: Comparison of seismic traces at the area of interest on TAMU Massif with the seismograms generated from best impedance model.

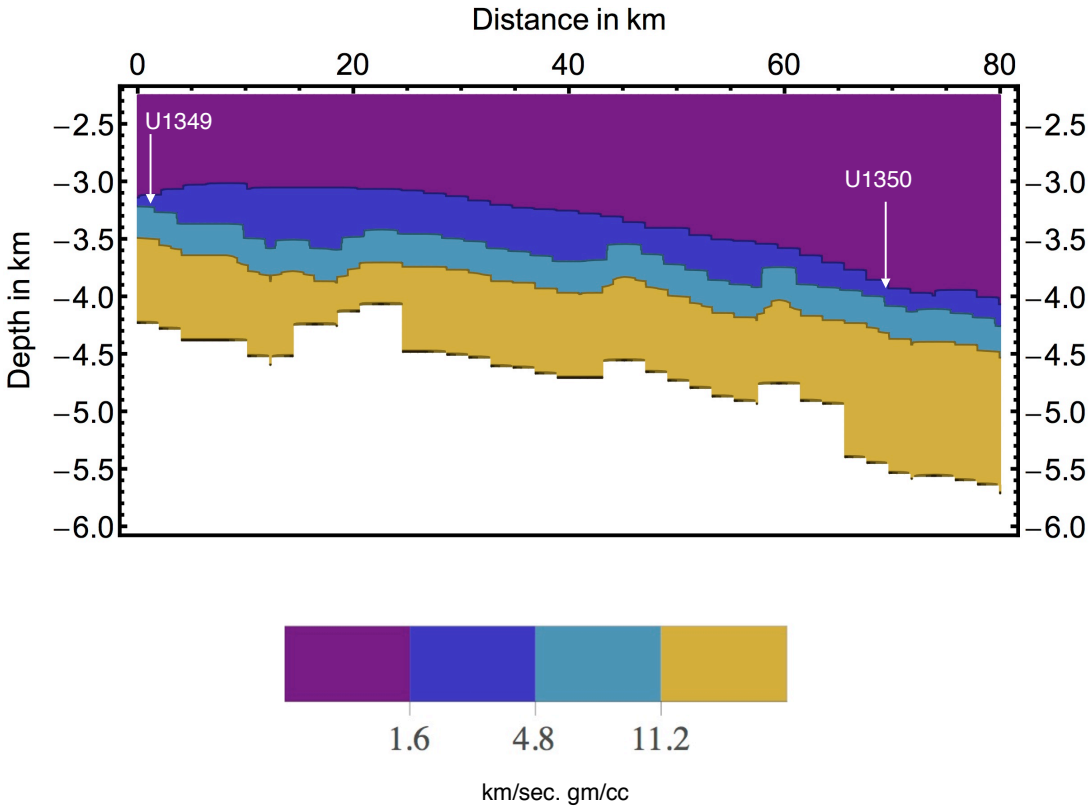


Figure 4.24: The prior impedance model in the area highlighted on Ori Massif in in Figure 4.3. The prior impedance model on Ori Massif is obtained by modifying the velocity profiles generated at TAMU Massif by Korenaga and Sager (2012).

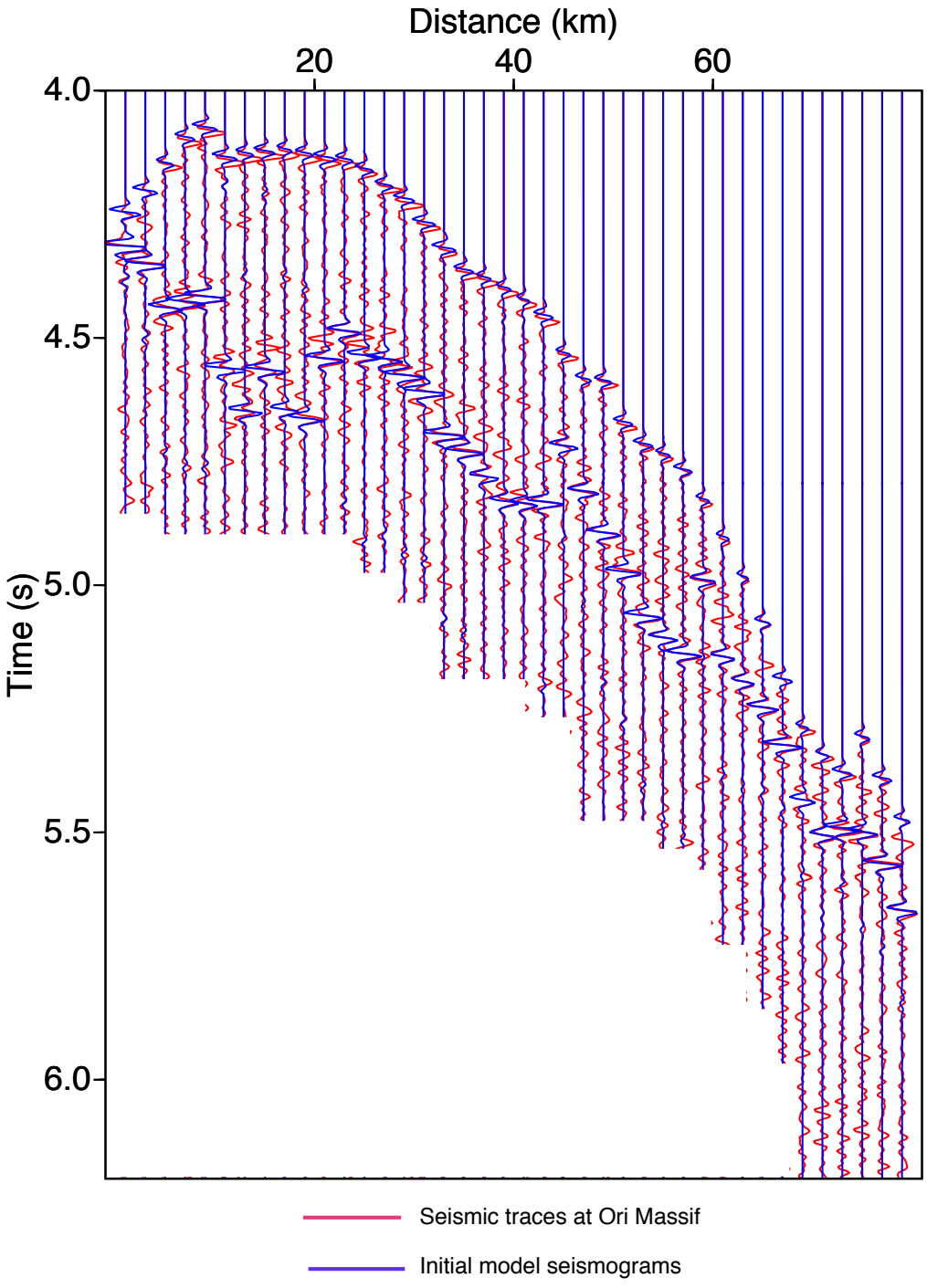


Figure 4.25: Comparison of seismic traces at the area of interest on Ori Massif with the seismograms generated from prior impedance model. The distance between seismic traces is 2 km. The location of the seismic section chosen here is shown as a rectangle on Ori Massif in Figure 4.3.

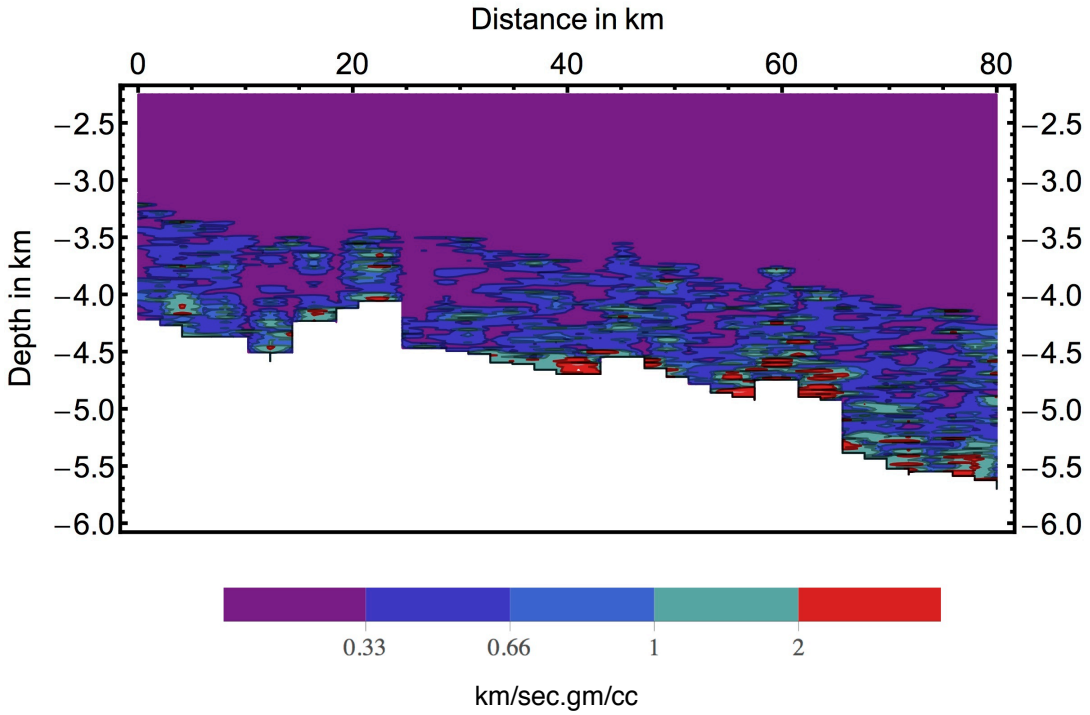


Figure 4.26: Standard deviation of impedance models in the area of interest on the Ori Massif.

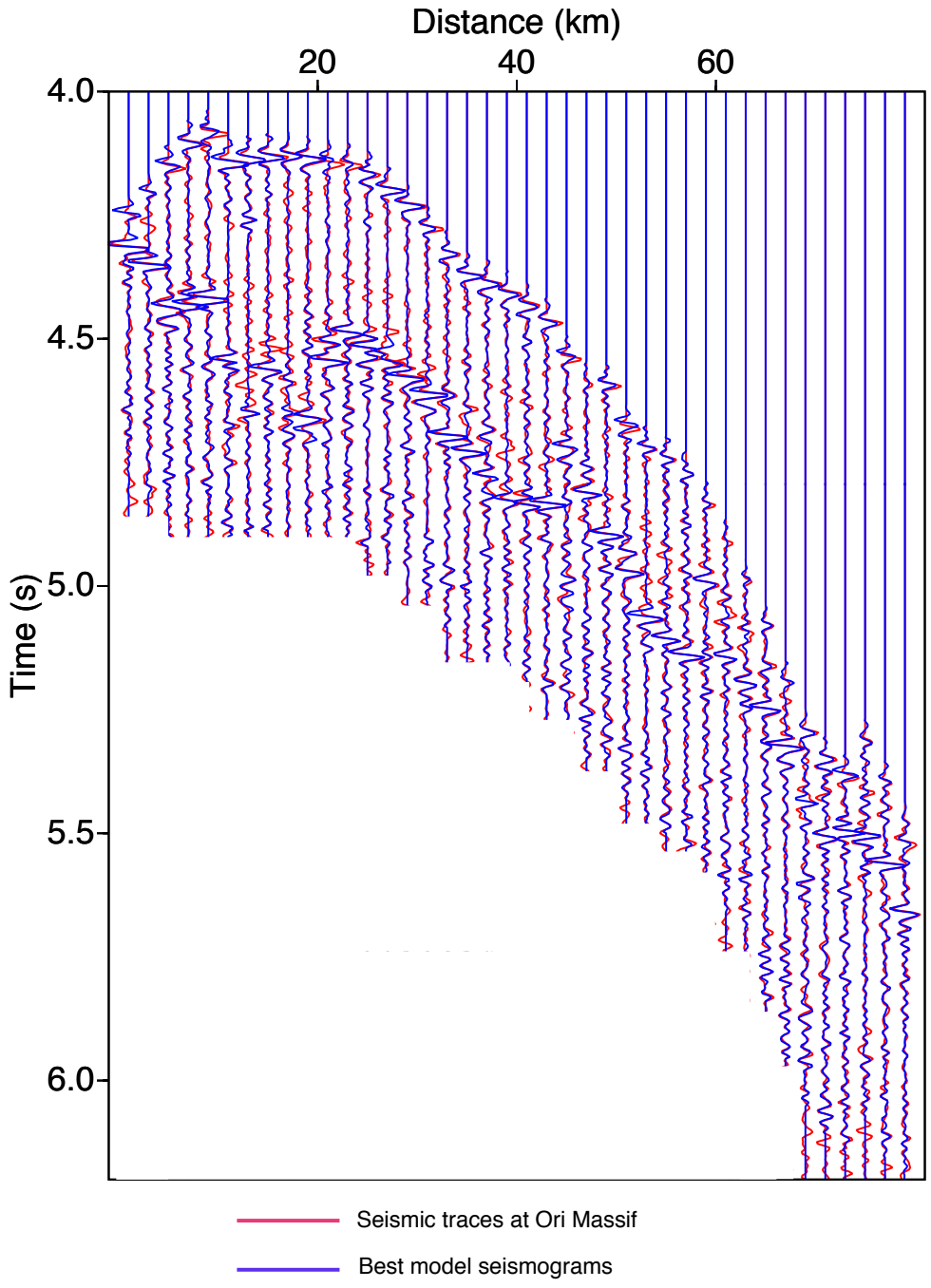
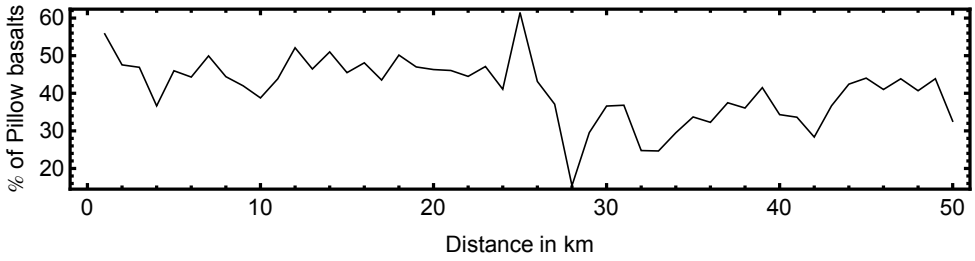
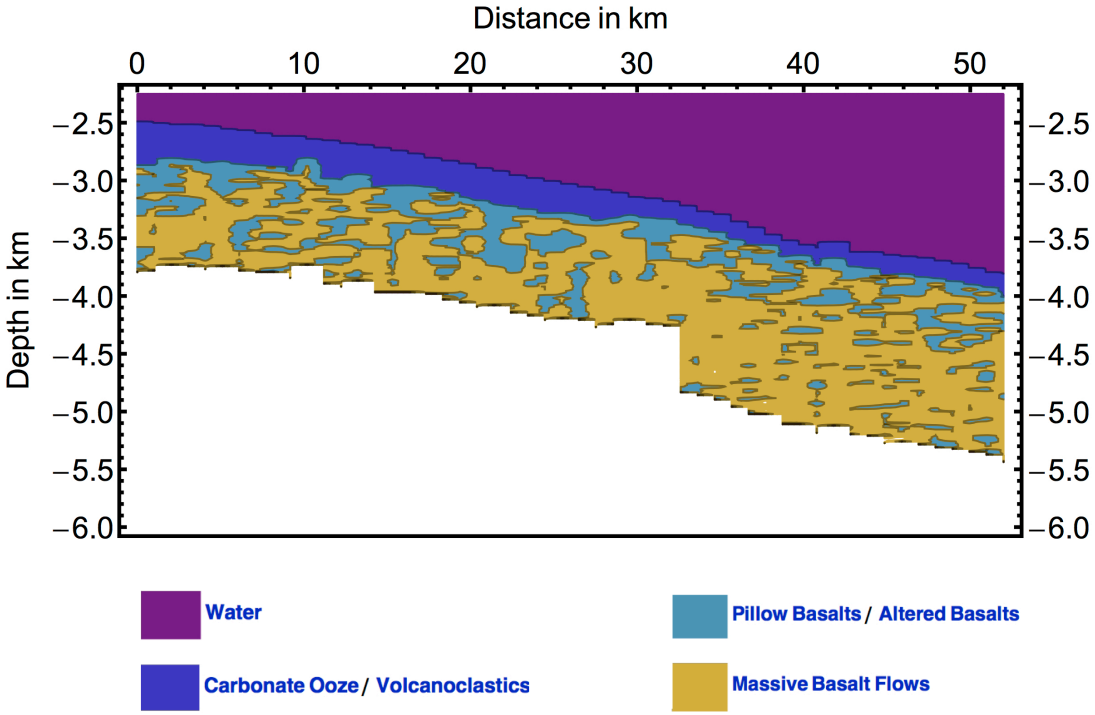


Figure 4.27: Comparison of seismic traces at the area of interest on TAMU Massif with the seismograms generated from best impedance model.



(a)



(b)

Figure 4.28: (a) A 1-D plot showing the variation of percentage of pillow basalts with distance at TAMU Massif. (b) Lithology in the area of interest on TAMU Massif classified based on a cut off value of 11.75 km/sec.gm/cc or above as massive flow basalts.

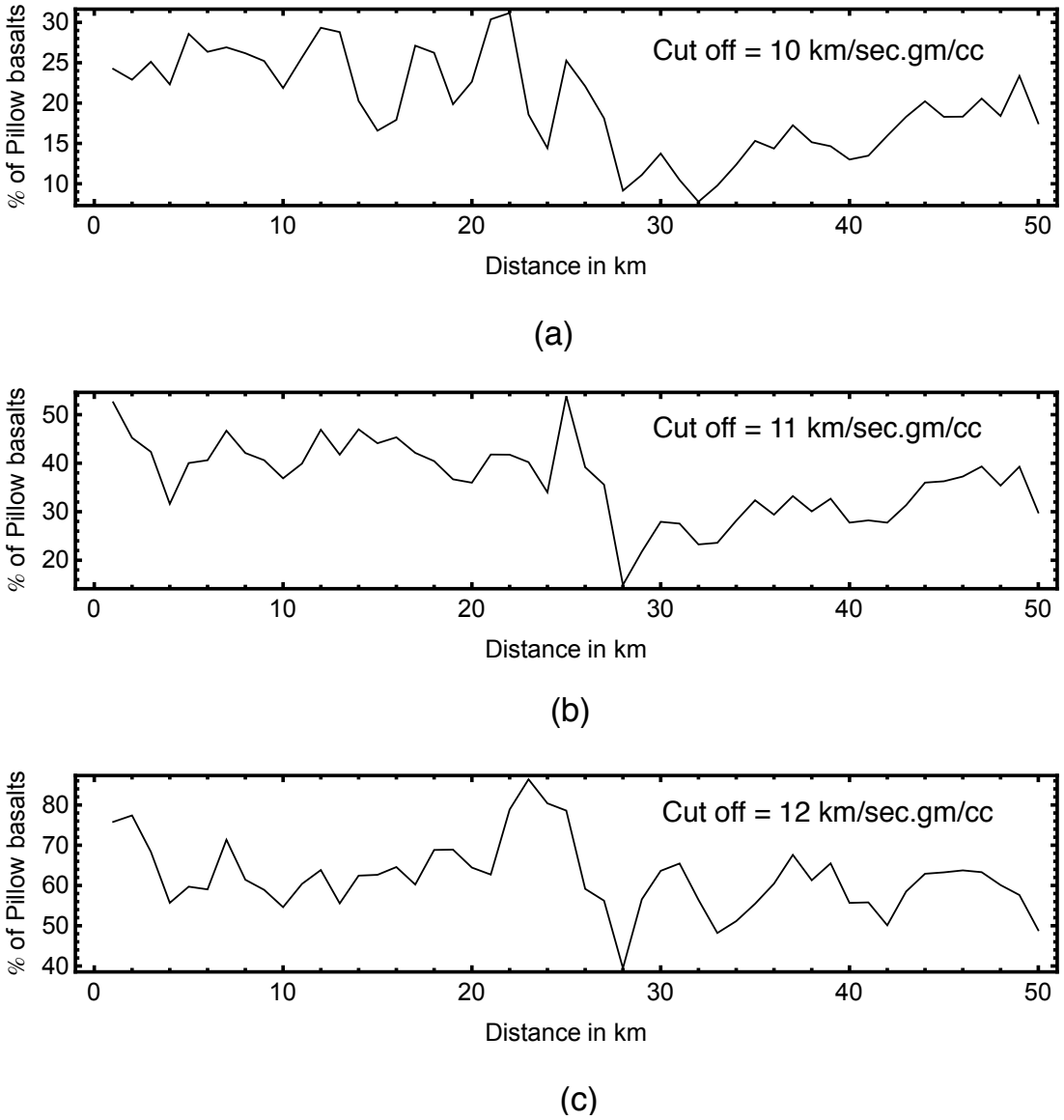
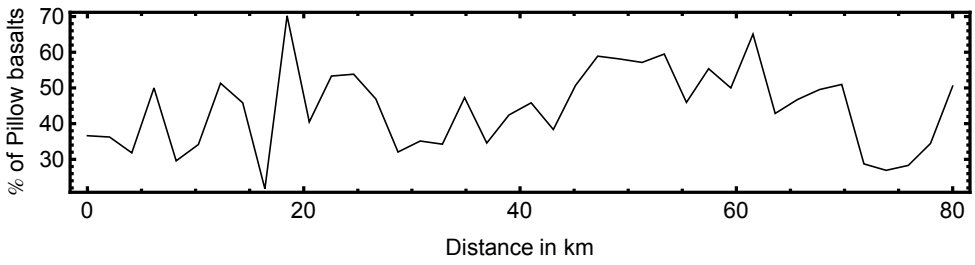
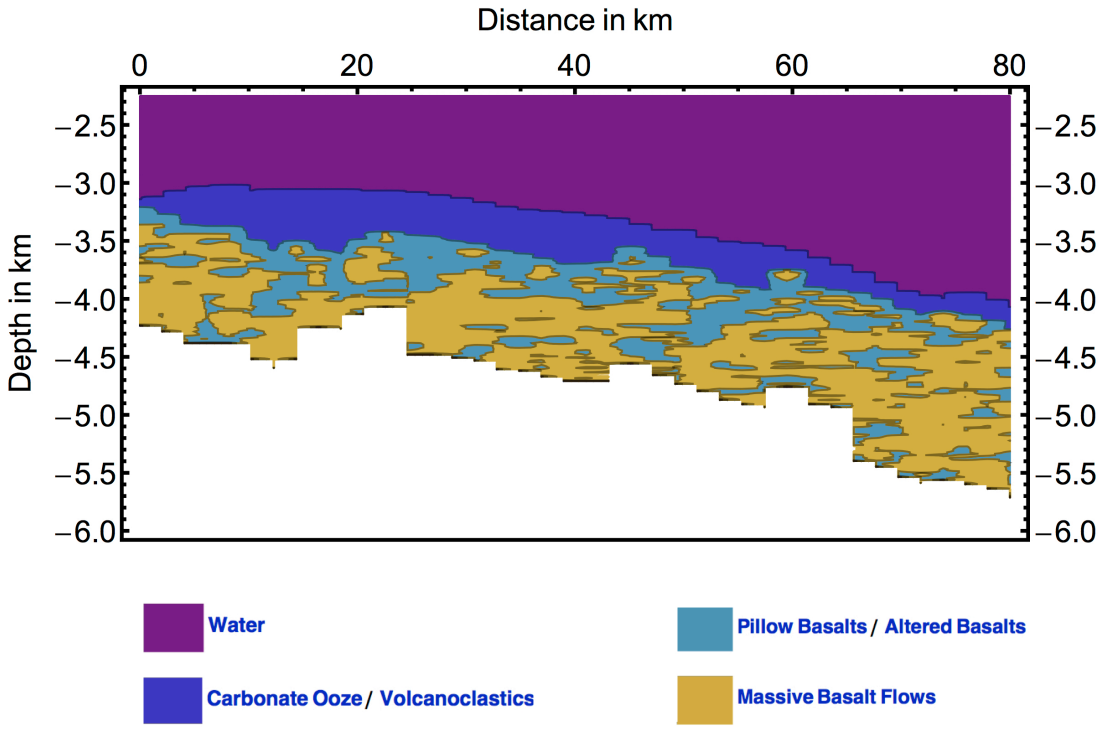


Figure 4.29: 1-D plots showing the variation of percentage of pillow basalts at TAMU Massif when the cut off for pillow/altered basalts is (a) 10 km/sec.gm/cc, (b) 11 km/sec.gm/cc and (c) 12 km/sec.gm/cc.



(a)



(b)

Figure 4.30: (a) A 1-D plot showing the variation of percentage of pillow basalts with distance at Ori Massif. (b) Lithology in the area of interest on Ori Massif classified based on a cut off value of 11.75 or above as massive flow basalts.

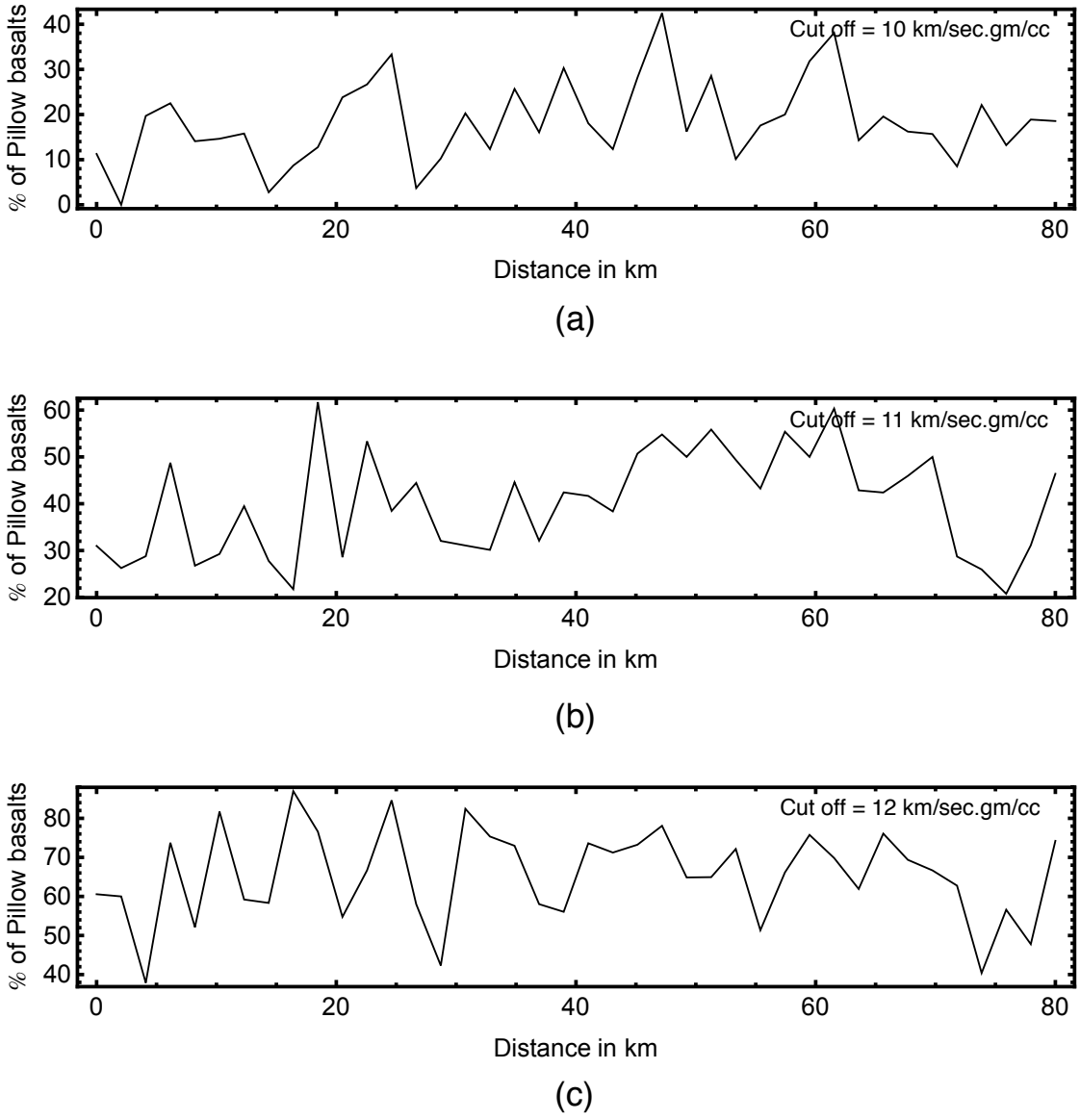


Figure 4.31: 1-D plots showing the variation of percentage of pillow basalts with distance at Ori Massif when the cut off for pillow/altered basalts is (a) 10 km/sec.gm/cc, (b) 11 km/sec.gm/cc and (c) 12 km/sec.gm/cc.

5. CONCLUSIONS

The dissertation uses a stochastic inversion technique for two applications: well log upscaling and seismic inversion. The stochastic inversion technique uses reversible jump MCMC sampling in a Bayesian framework. The model parameters during the random sweep include number of layers, layer boundaries and impedance in layers. In the first application, upscaling velocity logs is defined as an inversion problem where the thickness of the effective layer thicknesses of the optimal models depends on the wavelength of the propagating wave. Results on velocity log upscaling suggested the number of effective medium layers increases as the frequency of the reference seismogram increases. The effective medium velocity is calculated by adding a random perturbation to a deterministic Backus average value. I calculated the optimal models and uncertainty in velocity log upscaling using rjMCMC sampling with and without using Simulated Annealing respectively. The upscaled velocity when using normal incidence seismogram as reference seismogram is smoother compared to the case when performing log blocking with multi-offset seismograms as reference seismograms.

The second application is to obtain shallow impedance structure on TAMU and Ori massifs at the Shatsky Rise oceanic plateau. I compared the inverted impedance results with impedance from well log at U1347 and with lithology from drill cores at three location on Shatsky Rise for calibration study. Since impedance is a rock property, it is used to discriminate lithologies in shallow structure of TAMU and Ori massifs. The inferred lithology suggests TAMU massif has more massive flow basalts compared to Ori massif suggesting massive wide spread lava flows waned during the formation of Ori massif. The volume of pillow/altered basalts at the summit of the volcanoes is higher compared to that at the flank. The abundance

of lower impedance basalts at the summit compared to the flank can be due to localized explosive volcanism at the summit or due to geochemical alteration caused by fluid-rock interaction.

REFERENCES

Agostinetti, N. P., and A. Malinverno, 2010, Receiver function inversion by trans-dimensional Monte Carlo sampling: Geophysical Journal International, **181**, 858--872.

Aki, K., and P. G. Richards, 2002, Quantitative seismology: Theory and methods: University Science Books, Herndon, Virginia, USA.

Andrieu, C., N. De Freitas, and A. Doucet, 2000, Reversible jump MCMC simulated annealing for neural networks: Proceedings of the Sixteenth conference on Uncertainty in artificial intelligence, Morgan Kaufmann Publishers Inc., 11--18.

Avseth, P., T. Mukerji, and G. Mavko, 2005, Quantitative seismic interpretation: Cambridge University Press, Cambridge, United Kingdom.

Bartetzko, A., R. Pechnig, and J. Wohlenberg, 2002, Interpretation of well-logging data to study lateral variations in young oceanic crust: DSDP/ODP Holes 504B and 896A, Costa Rica Rift: Geological Applications of Well Logs, AAPG Methods in Exploration Series, **13**, 213--228.

Bodin, T., and M. Sambridge, 2009, Seismic tomography with the reversible jump algorithm: Geophysical Journal International, **178**, 1411--1436.

Bodin, T., M. Sambridge, H. Tkalćić, P. Arroucau, K. Gallagher, and N. Rawlinson, 2012, Transdimensional inversion of receiver functions and surface wave dispersion: Journal of Geophysical Research: Solid Earth (1978--2012), **117**.

Brooks, S. P., and A. Gelman, 1998, General methods for monitoring convergence of iterative simulations: Journal of Computational and Graphical statistics, **7**, 434--455.

Buland, A., and Y. El Ouair, 2006, Bayesian time-lapse inversion: Geophysics, **71**,

R43--R48.

Buland, A., and H. Omre, 2003, Bayesian linearized AVO inversion: Geophysics, **68**, 185--198.

Cardiff, M., and P. Kitanidis, 2009, Bayesian inversion for facies detection: An extensible level set framework: Water Resources Research, **45**.

Castelloe, J. M., and D. L. Zimmerman, 2002, Convergence assessment for reversible jump MCMC samplers: Department of Statistics and Actuarial Science, University of Iowa, Technical Report, **313**.

Coffin, M. F., and O. Eldholm, 1994, Large igneous provinces: crustal structure, dimensions, and external consequences: Reviews of Geophysics, **32**, 1--36.

De, G. S., D. F. Winterstein, and M. A. Meadows, 1994, Comparison of P-and S-wave velocities and Q's from VSP and sonic log data: Geophysics, **59**, 1512--1529.

Delacour, A., and D. Guillaume, 2013, Data report: alteration of basalts from Sites U1346 and U1349 at Shatsky Rise oceanic plateau, IODP Expedition 324: Proc. IODP, **324**, Chapter 203.

Denison, D., N. Adams, C. Holmes, and D. Hand, 2002, Bayesian partition modelling: Computational statistics & data analysis, **38**, 475--485.

Dettmer, J., C. W. Holland, and S. E. Dosso, 2013, Transdimensional uncertainty estimation for dispersive seabed sediments: Geophysics, **78**, WB63--WB76.

Duncan, R. A., and M. Richards, 1991, Hotspots, mantle plumes, flood basalts, and true polar wander: Reviews of Geophysics, **29**, 31--50.

Eidsvik, J., P. Avseth, H. Omre, T. Mukerji, and G. Mavko, 2004, Stochastic reservoir characterization using prestack seismic data: Geophysics, **69**, 978--993.

Folstad, P. G., and M. Schoenberg, 1992, *in* Low frequency propagation through fine layering: 62nd Annual International Meeting, SEG, Expanded Abstracts, 1279--1281.

Foulger, G. R., 2007, The‘‘plate’’model for the genesis of melting anomalies: Special papers-Geological Society of America, **430**, 1.

Gallagher, K., T. Bodin, M. Sambridge, D. Weiss, M. Kylander, and D. Large, 2011, Inference of abrupt changes in noisy geochemical records using transdimensional changepoint models: Earth and Planetary Science Letters, **311**, 182--194.

Gelman, A., and D. B. Rubin, 1992, Inference from iterative simulation using multiple sequences: Statistical science, **7**, 457--472.

Gibson, R. L., and K. L. Hwang, 2009, Quantification of uncertainty in velocity log upscaling by a Markov Chain Monte Carlo method: 2587--2591.

Gibson Jr, R. L., 2005, Influence of internal reservoir structure on composite reflection coefficients: **24**, 312--315.

Goetz, J., L. Dupal, and J. Bowler, 1979, An investigation into discrepancies between sonic log and seismic check shot velocities: Australian Petroleum Exploration Association Journal, **19**, 131--141.

Grana, D., and E. Della Rossa, 2010, Probabilistic petrophysical-properties estimation integrating statistical rock physics with seismic inversion: Geophysics, **75**, O21--O37.

Grechka, V., 2003, Effective media: A forward modeling view: Geophysics, **68**, 2055--2062.

Green, P. J., 1995, Reversible jump Markov Chain Monte Carlo computation and Bayesian model determination: Biometrika, **82**, 711--732.

--------, 2003, Trans-dimensional Markov Chain Monte Carlo: Oxford Statistical Science Series, 179--198.

Gunning, J., and M. E. Glinsky, 2004, Delivery: an open-source model-based bayesian seismic inversion program: Computers & Geosciences, **30**, 619--636.

Hastings, W. K., 1970, Monte Carlo sampling methods using Markov Chains and

their applications: Biometrika, **57**, 97--109.

Hong, T., and M. K. Sen, 2009, A new MCMC algorithm for seismic waveform inversion and corresponding uncertainty analysis: Geophysical Journal International, **177**, 14--32.

Ingber, L., 1989, Very fast simulated re-annealing: Mathematical and computer modelling, **12**, 967--973.

Ingle, S., and M. F. Coffin, 2004, Impact origin for the greater ontong java plateau?: Earth and Planetary Science Letters, **218**, 123--134.

JafarGandomi, A., and A. Binley, 2013, A Bayesian trans-dimensional approach for the fusion of multiple geophysical datasets: Journal of Applied Geophysics, **96**, 38--54.

Koppers, A., T. Sano, J. Natland, M. Widdowson, R. Almeev, A. Greene, and D. Murphy, 2010, Massive basalt flows on the southern flank of Tamu Massif, Shatsky Rise: a reappraisal of ODP Site 1213 basement units1: Proc. IODP, **324**, Chapter 109.

Korenaga, J., and W. Sager, 2012, Seismic tomography of Shatsky Rise by adaptive importance sampling: Journal of Geophysical Research: Solid Earth (1978--2012), **117**.

Lindsay, R., and R. Van Koughnet, 2001, Sequential Backus Averaging: Upscaling well logs to seismic wavelengths: The Leading Edge, **20**, 188--191.

Liner, C. L., and T. W. Fei, 2006, Layer-induced seismic anisotropy from full-wave sonic logs: Theory, application, and validation: Geophysics, **71**, D183--D190.

Mahoney, J., R. Duncan, M. Tejada, W. Sager, and T. Bralower, 2005, Jurassic-Cretaceous boundary age and mid-ocean-ridge--type mantle source for Shatsky Rise: Geology, **33**, 185--188.

Mahoney, J., M. Storey, R. Duncan, K. Spencer, and M. Pringle, 1993, Geochemistry

and age of the Ontong Java Plateau: Geophysical Monograph Series, **77**, 233--261.

Malinverno, A., 2002, Parsimonious Bayesian Markov Chain Monte Carlo inversion in a nonlinear geophysical problem: Geophysical Journal International, **151**, 675--688.

Malinverno, A., and W. S. Leaney, 2005, Monte-Carlo Bayesian look-ahead inversion of walkaway vertical seismic profiles: Geophysical prospecting, **53**, 689--703.

Mondal, A., Y. Efendiev, B. Mallick, and A. Datta-Gupta, 2010, Bayesian uncertainty quantification for flows in heterogeneous porous media using reversible jump Markov Chain Monte Carlo methods: Advances in Water Resources, **33**, 241--256.

Mosegaard, K., and A. Tarantola, 1995, Monte Carlo sampling of solutions to inverse problems: Journal of Geophysical Research: Solid Earth (1978--2012), **100**, 12431--12447.

Prüssmann, J., 1996, Better preservation of seismic fine layering effects by new methods for blocking log data: Presented at the 58th EAGE Meeting in Amsterdam, The Netherlands.

Ramirez, A., D. White, Y. Hao, K. Dyer, and J. Johnson, 2013, Estimating reservoir permeabilities using the seismic response to CO2 injection and stochastic inversion: International Journal of Greenhouse Gas Control, **16, Supplement 1**, S146 -- S159.

Reading, A. M., and K. Gallagher, 2013, Transdimensional change-point modeling as a tool to investigate uncertainty in applied geophysical inference: An example using borehole geophysical logs: Geophysics, **78**, WB89--WB99.

Richards, M. A., R. A. Duncan, and V. E. Courtillot, 1989, Flood basalts and hot-spot tracks: Plume heads and tails: Science, **246**, 103--107.

Rimstad, K., P. Avseth, and H. Omre, 2010, Bayesian lithology/fluid prediction constrained by spatial couplings and rock physics depth trends: The Leading Edge, **29**, 584--589.

Rimstad, K., and H. Omre, 2010, Impact of rock-physics depth trends and Markov random fields on hierarchical Bayesian lithology/fluid prediction: Geophysics, **75**, R93--R108.

Rio, P., T. Mukerji, G. Mavko, and D. Marion, 1996, Velocity dispersion and upscaling in a laboratory-simulated VSP: Geophysics, **61**, 584--593.

Sager, W., and H.-C. Han, 1993, Rapid formation of the Shatsky Rise oceanic plateau inferred from its magnetic anomaly: Nature, **364**, 610--613.

Sager, W. W., T. Sano, J. Geldmacher, et al., 2011, IODP Expedition 324: Ocean Drilling at Shatsky Rise Gives Clues about Oceanic Plateau Formation: Scientific Drilling, **12**, 24--31.

Sager, W. W., J. Zhang, J. Korenaga, T. Sano, A. A. Koppers, M. Widdowson, and J. J. Mahoney, 2013, An immense shield volcano within the Shatsky Rise oceanic plateau, northwest Pacific Ocean: Nature Geoscience, **6**, 976--981.

Sambridge, M., 1999, Geophysical inversion with a neighbourhood algorithm---I. Searching a parameter space: Geophysical Journal International, **138**, 479--494.

Sambridge, M., T. Bodin, K. Gallagher, and H. Tkalčić, 2013, Transdimensional inference in the geosciences: Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences, **371**.

Sano, T., K. Shimizu, A. Ishikawa, R. Senda, Q. Chang, J.-I. Kimura, M. Widdowson, and W. W. Sager, 2012, Variety and origin of magmas on Shatsky Rise, northwest Pacific Ocean: Geochemistry, Geophysics, Geosystems, **13**.

Sayers, C., 1998, Long-wave seismic anisotropy of heterogeneous reservoirs: Geophysical Journal International, **132**, 667--673.

Sen, M. K., and P. L. Stoffa, 1991, Nonlinear one-dimensional seismic waveform inversion using simulated annealing: Geophysics, **56**, 1624--1638.

Tiwary, D. K., I. O. Bayuk, A. Vikhorev, M. Ammerman, and E. M. Chesnokov,

2007, Comparison of seismic upscaling methods: 1933--1936.

Ulrych, T. J., M. D. Sacchi, and A. Woodbury, 2001, A Bayes tour of inversion: A tutorial: Geophysics, **66**, 55--69.

Ulvmoen, M., and H. Omre, 2010, Improved resolution in Bayesian lithology/fluid inversion from prestack seismic data and well observations: Part 1---Methodology: Geophysics, **75**, R21--R35.

Zhang, J., W. W. Sager, and J. Korenaga, 2012, Shatsky Rise Oceanic Plateau Structure from 2D Multichannel Seismic Reflection Profiles and Implications for Oceanic Plateau Evolution: Eos. DI53A-2360 presented at 2012 AGU Fall Meeting in San Francisco.

APPENDIX A

ACCEPTANCE CRITERIA

In this appendix, I derive the Metropolis-Hastings (M-H) ratio for rjMCMC sampling which defines the criteria to accept and reject samples. I start with deriving the prior and proposal ratios which are necessary to calculate the M-H ratio.

A.1 Prior ratio

The prior probability distribution for the model parameters is given by

$$p(\mathbf{m}) = p(k)p(z, v/k), \tag{A.1}$$

where $\mathbf{m}$ is a vector (k, z, v). The prior conditional distribution of depth of interfaces (z) and layer velocities (v) given the number of layers (k) are considered as independent distributions. Although the prior conditional distributions for the above mentioned model parameters are independent, the posterior distributions might have a correlation. A physical explanation for this correlation can be explained by travel time being a function of velocity and layer thickness. A layer with an arbitrary velocity will result in similar travel time as a thiner layer with higher velocity which would result in an negative correlation between posterior distributions of velocity and thickness of layers. Significant deviations of perturbed velocities from the deterministic Backus average are not expected; the prior bounds are assigned to the layer velocities as follows,

$$p(v_i/k) = \begin{cases} 1/\Delta v & \text{if } V_{min} \leq v_i \leq V_{max} \\ 0 & \text{otherwise,} \end{cases} \tag{A.2}$$

where Δv is $(V_{max} - V_{min})$. In this study, $\Delta v = 1$ for all the examples.

The prior velocity distributions for all the layers are assumed as independent and identical distributions, as a result of which the joint distribution of velocities can be written as product of velocity distributions for individual layers. To define $p(z/k)$ we followed the approach discussed in the work of Bodin and Sambridge (2009) and Bodin et al. (2012). The interval of interest can be divided into N cells and the k boundaries can be placed in $\frac{N!}{k!(N-k)!}$ possible ways. The probability of occurrence of each model, $p(z/k)$, is $(\frac{N!}{k!(N-k)!})^{-1}$. Substituting equation (A.2) and $p(z/k)$ into equation (A.1) results in

$$p(m) = \frac{k!(N-k)!}{N!\Delta k(\Delta v)^k}, \tag{A.3}$$

where Δk defines the range of number of layers allowed.

Prior ratio for boundary change move $p(m')/p(m) = 1$ as the number of layers remain same in a boundary change move.

Prior ratio for birth step is given by

$$\frac{p(m')}{p(m)} = \frac{k+1}{(N-k)\Delta v}. \tag{A.4}$$

Prior ratio for death step is given by

$$\frac{p(m')}{p(m)} = \frac{(N-k+1)\Delta v}{k}. \tag{A.5}$$

A.2 Proposal ratio

The acceptance term which is the criteria for accepting models in a Markov Chain, is the product of prior ratio, proposal ratio and the likelihood ratio. In the Markov Chain, model parameter m' in iteration $i+1$ is chosen conditional only to

m in iteration i based on a proposal distribution. The convergence of the Markov Chain to a stationary distribution depends on the parameters chosen in the proposal distribution. The velocity of a randomly selected layer is perturbed in all the three moves based on the proposal distribution given by

$$q(\frac{v'}{m}) = q(\frac{v'}{v}) = \frac{1}{\theta\sqrt{2\Pi}} \exp\{\frac{-(v'-v)^2}{2\theta^2}\}. \tag{A.6}$$

In the boundary change move when no change of dimension is involved, the proposal distributions for boundary depths and velocities are symmetrical. So $q(z'/z) = q(z/z')$, $q(v'/v) = q(v/v')$ and the ratio of proposal distributions for number of layers (i.e $q(k'/k)$ and $q(k/k')$) is one, hence the proposal ratio $\frac{q(m'/m)}{q(m/m')} = 1$.

Birth step proposal ratio is

$$\frac{q(m/m')}{q(m'/m)} = \frac{N-k}{(k+1)q(v'/v)}. \tag{A.7}$$

Death step proposal ratio is

$$\frac{q(m/m')}{q(m'/m)} = \frac{kq(v'/v)}{N-k+1}. \tag{A.8}$$

Refer to Appendix B and Appendix C1 in Bodin et al. (2012) for detailed derivation of proposal ratio's for birth and death moves.

A.3 Acceptance term

A.3.1 Boundary move

$$\alpha(\mathbf{m}', \mathbf{m}) = min[1, \exp\{-\frac{\Phi(m') - \Phi(m)}{2}\}] \tag{A.9}$$

where $\Phi(m)$ is the likelihood function for the model parameter m given by equation

(2.2). The prior ratio and proposal ratio for boundary move are equal to 1 as explained in previous sections in the Appendix.

A.3.2 Birth move

$$\alpha(\mathbf{m}', \mathbf{m}) = min[1, \frac{\theta\sqrt{2\Pi}}{\Delta v} \exp\{\frac{(v_i' - v_i)^2}{2\theta^2} - \frac{\Phi(m') - \Phi(m)}{2}\}] \tag{A.10}$$

Equation (A.10) is obtained by substituting equations (A.4) and (A.7) in equation (2.4)

A.3.3 Death move

$$\alpha(\mathbf{m}', \mathbf{m}) = min[1, \frac{\Delta v}{\theta\sqrt{2\Pi}} \exp\{\frac{-(v_i' - v_i)^2}{2\theta^2} - \frac{\Phi(m') - \Phi(m)}{2}\}] \tag{A.11}$$

Equation (A.11) is obtained by substituting equations (A.5) and (A.8) in equation (2.4)

A.4 Parsimony

The acceptance term can be written as a product of two terms, the likelihood ratio and the Occam factor. The Occam factor is the product of prior ratio and the proposal ratio. The likelihood ratio increases with the number of layers as the match improves. The Occam factor during the birth move is

$$\frac{\theta\sqrt{2\Pi}}{\Delta v} \exp\{\frac{(v_i' - v_i)^2}{2\theta^2}\}, \tag{A.12}$$

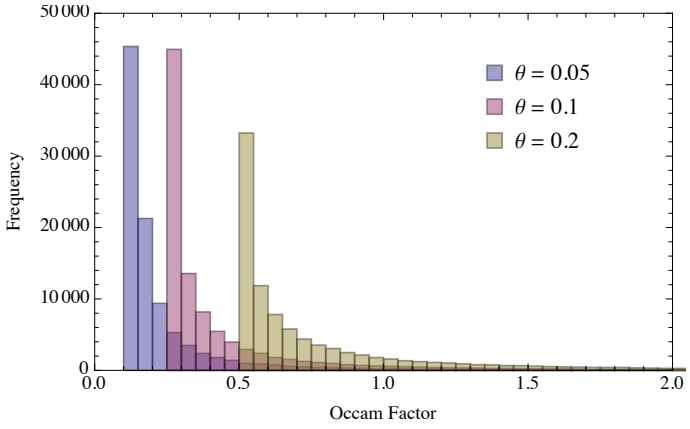
where $v_i' - v_i$ is the velocity perturbation from layer-based Backus average values. θ is the standard deviation of the velocity perturbation. The Occam factor during the

death move is

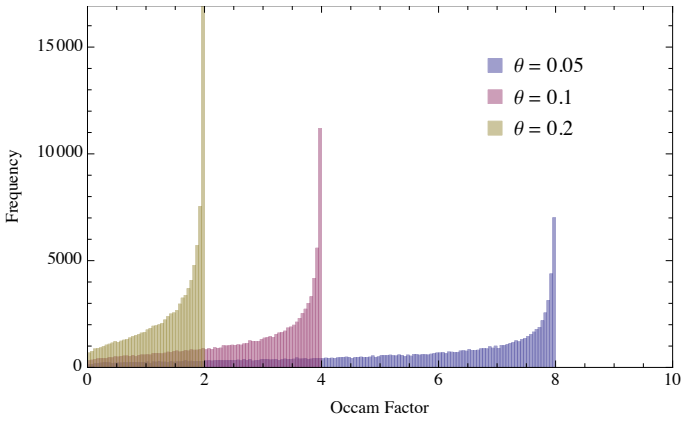
$$\frac{\Delta v}{\theta\sqrt{2\Pi}} \exp\{\frac{-(v_i' - v_i)^2}{2\theta^2}\}. \tag{A.13}$$

As observed in Figure A.1(a) the Occam factor is most likely less than 1 for any realistic standard deviation of proposal velocity perturbation during the birth step. The converse is true for the death step where the Occam factor is most likely greater than 1 (Figure A.1(b)).

When the layer number is small, the misfit term, and hence the likelihood ratio, dominates the acceptance probability. As the number of layers, increase, the model estimates the reference more accurately, the improvement from adding more layers become tangential, and hence the likelihood ratio is reduced. In this situation, the Occam factor starts to determine the acceptance ratio, which then will favor fewer layers as long as the model still can provide good enough prediction to the data. Through making the number of layers as another parameter in the posterior distribution modeling, we have naturally embedded the model selection problem into the Bayesian framework, and the rjMCMC simulated annealing is guaranteed to locate the optimal layer number.



(a)



(b)

Figure A.1: Occam factor histograms for birth and death moves.